

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2052	100.0	2052	6	E40087	E40087 Plant promo
2	2048.8	99.8	2052	6	E40093	E40093 Plant promo
3	2038	99.3	2056	6	E40091	E40091 Plant promo
4	2034	99.1	2048	6	E40090	E40090 Plant promo
5	2032.4	99.0	2048	6	E40089	E40089 Plant promo
6	241.2	11.8	2042	6	AR076817	AR076817 Sequence
7	241.2	11.8	2042	6	E15125	E15125 Promoter. 7
8	196.6	9.6	247	6	AR076816	AR076816 Sequence
9	196.6	9.6	247	6	E15124	E15124 Promoter. 7
10	195.6	9.5	246	6	AR146852	AR146852 Sequence
11	195.6	9.5	246	6	E55065	E55065 Plant promo
C	174	8.5	8005	8	DCA18706	Y18706 Daucus promo
13	162.4	7.9	2831	6	BD188672	BD188672 Promotor
14	162.4	7.9	2865	6	BD188674	BD188674 Promotor
15	162.4	7.9	4886	8	DARGCH52	D16255 Carrot gCHS
16	119	5.8	140	6	AR146856	AR146856 Sequence
C	119	5.8	140	6	AR146857	AR146857 Sequence
18	119	5.8	140	6	E55069	E55069 Plant promo
C	119	5.8	140	6	E55070	E55070 Plant promo

Db	1741	GTTTGAACAATGATGTCGGGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT	1800
Qy	1801	GCATTTCTAGAAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG	1860
Db	1801	GCATTTCTAGAAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG	1860
Qy	1861	GATTGGAATCCTTTTCTAAACTTTTAAATATAAAAAATGCAATATTGTAATATTTATC	1920
Db	1861	GATTGGAATCCTTTTCTAAACTTTTAAATATAAAAAATGCAATATTGTAATATTTATC	1920
Qy	1921	AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT	1980
Db	1921	AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT	1980
Qy	1981	CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATAAC	2040
Db	1981	CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATCAC	2040
Qy	2041	ATTTCTAAATATC 2052	
Db	2041	ATTTCTAAAGATC 2052	
RESULT 3			
E40091			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2052; Conservative			
Qy	1	CATGTGTGCCCTACAGCACATAGGGCCTGTTGGTTGGTGGAGAGCAGAAAGCTGCTCTGA	60
Db	1	CATGTGTGCCCTACAGCACATAGGGCCTGTTGGTTGGTGGAGAGCAGAAAGCTGCTCTGA	60
Qy	61	CTTCTTCTCTTTTGACCTGTTGTATTAAGAAGTAGAATAATTTTTAAAAAGCTGGAA	120
Db	61	CTTCTTCTCTTTTGACCTGTTGTATTAAGAAGTAGAATAATTTTTAAAAAGCTGGAA	120
Qy	121	TACTAACTTCTCTCAACCTCCGCTCTTTTCCAAACACTTTTAACTTTTCTTACT	180
Db	121	TACTAACTTCTCTCAACCTCCGCTCTTTTCCAAACACTTTTAACTTTTCTTACT	180

Qy	181	TTCTATTTCTACTCCACTTCTTTTGCTATTAAGCAAGAAATCACCTTCTTTTAAGCTAACCCA	240
Db	181	TTCTATTTCTACTCCACTTCTTTTGCTATTAAGCAAGAAATCACCTTCTTTTAAGCTAACCCA	240
Qy	241	AACGGCTCAATAAAGATCATTCAATAATGTAATCTTTCAATTTTAGGATAACAATACGT	300
Db	241	AACGGCTCAATAAAGATCATTCAATAATGTAATCTTTCAATTTTAGGATAACAATACGT	300
Qy	301	GAACAGGGTTATTTTTTAACGGTCAACAAATTTCTAATAATTTTACCTGCGCGGTGAACA	360
Db	301	GAACAGGGTTATTTTTTAACGGTCAACAAATTTCTAATAATTTTACCTGCGCGGTGAACA	360
Qy	361	CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAAACAAATTCGCATGC	420
Db	361	CCGTCTTCCAAGATAATATATTTTAATTTTGTAGCCTCCCTTTTAAACAAATTCGCATGC	420
Qy	421	AGGACGACTTGTAGTGAATACACATTTGCTGTGAGTCTTTTAAACAAAGAACAGTGGTTC	480
Db	421	AGGACGACTTGTAGTGAATACACATTTGCTGTGAGTCTTTTAAACAAAGAACAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAATTTGACAAAAACCCGACACACACCTCTATCCACGTACTATCTTT	540
Db	481	ATGCTCAGCCATCAAAATTTGACAAAAACCCGACACACACCTCTATCCACGTACTATCTTT	540
Qy	541	TGCGCGAATGCTTCTCAAAATGTTTTTATATGTATAAATAATGCCCATCCAGGATAAGT	600
Db	541	TGCGCGAATGCTTCTCAAAATGTTTTTATATGTATAAATAATGCCCATCCAGGATAAGT	600
Qy	601	AAATTTCCCGTTTAAACAGTTTCTTAATATATATGTTTACACATTACAAGAGGATATTCGT	660
Db	601	AAATTTCCCGTTTAAACAGTTTCTTAATATATATGTTTACACATTACAAGAGGATATTCGT	660
Qy	661	AATACTTTTGTAGCACAAGAGACTTAGGTCAAAATAATGGACGCTGGTAAACAGCCTAGACT	720
Db	661	AATACTTTTGTAGCACAAGAGACTTAGGTCAAAATAATGGACGCTGGTAAACAGCCTAGACT	720
Qy	721	TGCTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTAAG	780
Db	721	TGCTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTAAG	780
Qy	781	ATTAGAGCTATTAAATTAAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAACAGGTA	840
Db	781	ATTAGAGCTATTAAATTAAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAACAGGTA	840
Qy	841	AAAAACAAGAGCTTGTGCTGTGTTAGTTGTTGAGCTCATTTCTTTAAAAAGTAATG	900
Db	841	AAAAACAAGAGCTTGTGCTGTGTTAGTTGTTGAGCTCATTTCTTTAAAAAGTAATG	900
Qy	901	TAAACTGATCTAAAGCACATAGAAAAATTTAGTACAGGTTAAAACTTTTACAGAAATTTATA	960
Db	901	TAAACTGATCTAAAGCACATAGAAAAATTTAGTACAGGTTAAAACTTTTACAGAAATTTATA	960
Qy	961	TTAAACGAAATCATTTTAAACATGCTCTCGGCTGTCTAATTAATAGGATCACTTAC	1020
Db	961	TTAAACGAAATCATTTTAAACATGCTCTCGGCTGTCTAATTAATAGGATCACTTAC	1020
Qy	1021	TGATCATCTCAATTAACAACTTTGTTAAAAACAAATTTCAATGAGATAAAATATCTTACAATGAA	1080
Db	1021	TGATCATCTCAATTAACAACTTTGTTAAAAACAAATTTCAATGAGATAAAATATCTTACAATGAA	1080
Qy	1081	AAGAAGACAATGTCTCTTTGAAAAAACAATAAGGTAAGTCTCCCTCGCTCTGAAATGT	1140
Db	1081	AAGAAGACAATGTCTCTTTGAAAAAACAATAAGGTAAGTCTCCCTCGCTCTGAAATGT	1140
Qy	1141	ATACATATGATTGGACACGAGACTAAGAAAAATGATAAAGTAATGTAAGTAAAGGTAAG	1200
Db	1141	ATACATATGATTGGACACGAGACTAAGAAAAATGATAAAGTAATGTAAGTAAAGGTAAG	1200
Qy	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAACCAATATATAATTTGATAGATTAG	1260
Db	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAACCAATATATAATTTGATAGATTAG	1260

Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATATAAAATTTACTATTTTG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAATTTACTATTTTG 1320
Qy 1321 AGAAAGTTTTCGAAATGTATAGAAATGTAGTGGGACATCCATATAAAAGGAAAGTGTATAGAA 1380
Db 1321 AGAAAGTTTTCGAAATGTATAGAAATGTAGTGGGACATCCATATAAAAGGAAAGTGTATAGAA 1380
Qy 1381 TAAATGGGACAGAGGGAGTAACTCTTTATCATATATATAAATTTTGTATTTTGTATTTTCA 1440
Db 1381 TAAATGGGACAGAGGGAGTAACTCTTTATCATATATATAAATTTTGTATTTTGTATTTTCA 1440
Qy 1441 TAAGATTATAAATCTATGTGTTATTAATGATAATATAAATTTTAAAAATAAATACTATATAATT 1500
Db 1441 TAAGATTATAAATCTATGTGTTATTAATGATAATATAAATTTTAAAAATAAATACTATATAATT 1500
Qy 1501 CTGATTAGTCGATTTACCGCCTTTTATTAATTTTACAACTAGTAAATATGAAATTAATCAG 1560
Db 1501 CTGATTAGTCGATTTACCGCCTTTTATTAATTTTACAACTAGTAAATATGAAATTAATCAG 1560
Qy 1561 TTATCTGAAAAGCAAAATAATCTTTGTAAACAGCGTTCCGTCAAAATGGGAAGTTCATG 1620
Db 1561 TTATCTGAAAAGCAAAATAATCTTTGTAAACAGCGTTCCGTCAAAATGGGAAGTTCATG 1620
Qy 1621 TGTATTCGAATAGTTTAAATATAAAGTAAATTTTAAATTTTAAATTTTGTTCAGA 1680
Db 1621 TGTATTCGAATAGTTTAAATATAAAGTAAATTTTAAATTTTAAATTTTGTTCAGA 1680
Qy 1681 AATTAAAAATAATTTAGCATGGGAAGTTCACGGGCATCATTTGAGCAGCAGTACTAGACT 1740
Db 1681 AATTAAAAATAATTTAGCATGGGAAGTTCACGGGCATCATTTGAGCAGCAGTACTAGACT 1740
Qy 1741 GTTTGAAACAATGTATGTCGGGTGACATCTATGACCTTTCAAACTCAAACTAGTCAATAAT 1800
Db 1741 GTTTGAAACAATGTATGTCGGGTGACATCTATGACCTTTCAAACTCAAACTAGTCAATAAT 1800
Qy 1801 GCATT---CTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTCTTC 1856
Db 1801 GCATTCTAGCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTCTTC 1860
Qy 1857 AACGGATTGGATCCTTTTCTAACTTTTAAATATAAATAAATAAATGCAATTTATGTAATATT 1916
Db 1861 AACGGATTGGAACTCCTTTTCTAACTTTTAAATAAATAAATAAATGCAATTTATGTAATATT 1920
Qy 1917 TATCAACACCTTCAACATTTAGTGTAGGCTACTATAAATAGTGTCTTTGGTGTCTACTA 1976
Db 1921 TATCAACACCTTCAACATTTAGTGTAGGCTACTATAAATAGTGTCTTTGGTGTCTACTA 1980
Qy 1977 TCATCACAATCAATCTTACACCACAAACCTTGAGCTTAAATTTTCTACTTATTTCTCAGCAA 2036
Db 1981 TCATCACAATCAATCTTACACCACAAACCTTGAGCTTAAATTTTCTACTTATTTCTCAGCAA 2040
Qy 2037 TAAACATCTAAATATC 2052
Db 2041 TAAACATCTAAATATC 2056

RESULT 4
E40090
LOCUS E40090 2048 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40090
VERSION E40090.1 GI:18627206
KEYWORDS JP 2000166577-A/4.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2048)
AUTHORS Nishikawa, S. and Oeda, K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 4 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.

PN JP 2000166577-A/4
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR SATOMI NISHIKAWA, KENJI OEDA
PI C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC
PC C12N15/00,
C12N5/00, (C12N5/00, C12R1:91)
CC Key Location/Qualifiers
FH promoter Location/Qualifiers
FT 1. .2048
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES
source
ORIGIN
Query Match 99.1%; Score 2034; DB 6; Length 2048;
Best Local Similarity 99.8%; Pred. No. 2.8e-307;
Matches 2048; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 CATGTGTGCCCTACAGCACATAGGGCTGTGTTGGTTGAGAGAGCAGAGCTGCTCTGA 60
Db 1 CATGTGTGCCCTACAGCACATAGGGCTGTGTTGGTTGAGAGAGCAGAGCTGCTCTGA 60
Qy 61 CTTCTCTCTCTTTTGACCTGTTGTATAAAGAGTGAATAATTTTTTAAAAAGCTGCGAA 120
Db 61 CTTCTCTCTCTTTTGACCTGTTGTATAAAGAGTGAATAATTTTTTAAAAAGCTGCGAA 120
Qy 121 TACTAACTCTCTCTCAACAACCTTCCGCTTCTTTTCCAAACACATTTATTAACCTTTTACT 180
Db 121 TACTAACTCTCTCTCAACAACCTTCCGCTTCTTTTCCAAACACATTTATTAACCTTTTACT 180
Qy 181 TCTCATTTCTACCTCACTCTTTGCTATAAGCAAGAAATCACTCTCTTTTAAAGCTAACCCA 240
Db 181 TCTCATTTCTACCTCACTCTTTGCTATAAGCAAGAAATCACTCTCTTTTAAAGCTAACCCA 240
Qy 241 AACGGCTCAATAAAGATCAATCATAAATGTATCTTTCAATTTTAGATTAACAACTAGT 300
Db 241 AACGGCTCAATAAAGATCAATCATAAATGTATCTTTCAATTTTAGATTAACAACTAGT 300
Qy 301 GAACAGGGTTTATTTTAAAGCTGTCAAACAAATTTCTAAATAATTTTACCTGGCGGTGAACA 360
Db 301 GAACAGGGTTTATTTTAAAGCTGTCAAACAAATTTCTAAATAATTTTACCTGGCGGTGAACA 360
Qy 361 CCGTCTTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGATGC 420
Db 361 CCGTCTTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGATGC 420
Qy 421 AGGACGACTTGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGAAACAGTGGTTC 480
Db 417 AGGACGACTTGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGAAACAGTGGTTC 476
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGTACTATATCTTT 540
Db 477 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGTACTATATCTTT 536
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATATATATATATATATATATATATATATATAT 600
Db 537 TGGCCGAATGCTTCTCAAAATGTTTTTATATATATATATATATATATATATATATATATAT 596
Qy 601 AAAATTCCTGTTTAAACAGTCTTTTATATATATATATATATATATATATATATATATATAT 660
Db 597 AAAATTCCTGTTTAAACAGTCTTTTATATATATATATATATATATATATATATATATAT 656
Qy 661 AATACATTTTAGACGACAAAGAGACTTAGGTCAAAATTTGAGCTGCTGCTGCTGCTGCTGCT 720
Db 657 AATACATTTTAGACGACAAAGAGACTTAGGTCAAAATTTGAGCTGCTGCTGCTGCTGCTGCT 716
Qy 721 TGGTCACTGATAAATAGATAATTTGTTAGTATATATATATATATATATATATATATATATAT 780
Db 717 TGGTCACTGATAAATAGATAATTTGTTAGTATATATATATATATATATATATATATATATAT 776

```
Qy 781 ATTAGAGCTATTAAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Db      |||
Qy 777 ATTAGAGCTATTAAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 836
Db      |||
Qy 841 AAAACAGAGCTGCTGCTGTGTGTTAGTTGTGTGAGCTCATTTCTTTTAAAGTAATG 900
Db      |||
Qy 837 AAAACAGAGCTGCTGCTGTGTGTTAGTTGTGTGAGCTCATTTCTTTTAAAGTAATG 896
Db      |||
Qy 901 TAAACTGATCTAAAGCAGACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAATTTATA 960
Db      |||
Qy 897 TAAACTGATCTAAAGCAGACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAATTTATA 956
Db      |||
Qy 961 TTAACGAAATCAATTTTATAAACAATCTCTCGCTGTCTATATAATAGGATCAGTTAC 1020
Db      |||
Qy 957 TTAACGAAATCAATTTTATAAACAATCTCTCGCTGTCTATATAATAGGATCAGTTAC 1016
Db      |||
Qy 1021 TGATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATATCTTACAATGAA 1080
Db      |||
Qy 1017 TGATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATATCTTACAATGAA 1076
Db      |||
Qy 1081 AAGAAGGACAATGCTCTTTGAAAAAACAATAGGTACTCCCTCGTCCCTCTGAATGT 1140
Db      |||
Qy 1077 AAGAAGGACAATGCTCTTTGAAAAAACAATAGGTACTCCCTCGTCCCTCTGAATGT 1136
Db      |||
Qy 1141 ATACATATGGAATGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAG 1200
Db      |||
Qy 1137 ATACATATGGAATGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAG 1196
Db      |||
Qy 1201 AAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCACCAATATATAATTCATAGATTAG 1260
Db      |||
Qy 1197 AAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCACCAATATATAATTCATAGATTAG 1256
Db      |||
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTATATATATAAATAATTTACTATTTG 1320
Db      |||
Qy 1257 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTATATATAAATAATTTACTATTTG 1316
Db      |||
Qy 1321 AGAAAGTTTGAATGTATAGAAATGTAGTGGGACATCCATAAAGGAAAGTGTATAGAA 1380
Db      |||
Qy 1317 AGAAAGTTTGAATGTATAGAAATGTAGTGGGACATCCATAAAGGAAAGTGTATAGAA 1376
Db      |||
Qy 1381 TAAATGGACAGAGGAGTAACTTTATGATATATAAATTTTGTATTGTTGATTTC 1440
Db      |||
Qy 1377 TAAATGGACAGAGGAGTAACTTTATGATATATAAATTTTGTATTGTTGATTTC 1436
Db      |||
Qy 1441 TAAGATTATAAATCTATGTTATAATGATAATATATTTTAAAAATAACTATATTAAT 1500
Db      |||
Qy 1437 TAAGATTATAAATCTATGTTATAATGATAATATATTTTAAAAATAACTATATTAAT 1496
Db      |||
Qy 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATCTGAGTAATATGATAAATCAG 1560
Db      |||
Qy 1497 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATCTGAGTAATATGATAAATCAG 1556
Db      |||
Qy 1561 TTATCTGAAGAGCAATAATATCTTTGTAACACAGCGTTCGGTCAAAATGGAGGTCATG 1620
Db      |||
Qy 1557 TTATCTGAAGAGCAATAATATCTTTGTAACACAGCGTTCGGTCAAAATGGAGGTCATG 1616
Db      |||
Qy 1621 TGTATTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTTTCAGA 1680
Db      |||
Qy 1617 TGTATTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTTTCAGA 1676
Db      |||
Qy 1681 AATTAAATTAATATGAGCATGGGAGTTTCAGGGCATCATGTAGAGCAGCTAGACT 1740
Db      |||
Qy 1677 AATTAAATTAATATGAGCATGGGAGTTTCAGGGCATCATGTAGAGCAGCTAGACT 1736
Db      |||
Qy 1741 GTTTGAACAATGTATGTCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Db      |||
Qy 1737 GTTTGAACAATGTATGTCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796
Db      |||
Qy 1801 GCATTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG 1860
Db      |||
Qy 1797 GCATTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG 1856
Db      |||
```

```
Qy 1861 GATTGGAATCCTTTTCTAACTTTTAAATAAAAAAATGCAATTATTGTAATATTATC 1920
Db      |||
Qy 1857 GATTGGAATCCTTTTCTAACTTTTAAATAAAAAAATGCAATTATTGTAATATTATC 1916
Db      |||
Qy 1921 AACACCTCAACATTTGATTTAGCTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCAT 1980
Db      |||
Qy 1917 AACACCTCAACATTTGATTTAGCTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCAT 1976
Db      |||
Qy 1981 CACATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATAAC 2040
Db      |||
Qy 1977 CACATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATAAC 2036
Db      |||
Qy 2041 ATTCTAAATATC 2052
Db      |||
Qy 2037 ATTCTAAATATC 2048
Db      |||

RESULT 5
E40089
LOCUS      E40089          2048 bp          DNA          linear          PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION  E40089
VERSION    E40089.1  GI:18627205
KEYWORDS   JP 2000166577-A/3.
SOURCE     unidentified
           unclassified.
REFERENCE  1 (bases 1 to 2048)
AUTHORS   Nishikawa,S. and Oeda,K.
TITLE     Plant promoter and terminator
JOURNAL   Patent: JP 2000166577-A 3 20-JUN-2000;
          SUMITOMO CHEM CO LTD
COMMENT   OS Daucus carota L.
          PN JP 2000166577-A/3
          PD 20-JUN-2000
          PF 01-OCT-1999  JP 1999281475
          PR
          PI SATOMI NISHIKAWA, KENJI OEDA
          PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC
          C12N15/00,
          PC C12N5/00, (C12N5/00,C12R1:91)
          CC
          FH
          FT
          Key promoter Location/Qualifiers
          Location/Qualifiers
          1..2048
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"

FEATURES
             source
             Query Match 99.0%; Score 2032.4; DB 6; Length 2048;
             Best Local Similarity 99.8%; Pred. No. 5e-307;
             Matches 2047; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

ORIGIN
Qy 1 CATGTGTGCCCTACAGCACATAGGGCTGTTGGTTGAGAGAGCAGAGCTGCTCTGA 60
Db      |||
Qy 1 CATGTGTGCCCTACAGCACATAGGGCTGTTGGTTGAGAGAGCAGAGCTGCTCTGA 60
Db      |||
Qy 61 CTTCCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAAAGCTGGAA 120
Db      |||
Qy 61 CTTCCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAAAGCTGGAA 120
Db      |||
Qy 121 TACTAACTTCTCTCTCAAACTTCGCGTCTCTTTTCCAAACACTTTTAACTTTTACT 180
Db      |||
Qy 121 TACTAACTTCTCTCTCAAACTTCGCGTCTCTTTTCCAAACACTTTTAACTTTTACT 180
Db      |||
Qy 181 TCTCATTTCTACTCCACTTCTTGGCTATAGCAGAGAAATCACCTCTTTTAAAGCTAACCA 240
Db      |||
Qy 181 TCTCATTTCTACTCCACTTCTTGGCTATAGCAGAGAAATCACCTCTTTTAAAGCTAACCA 240
Db      |||
Qy 241 AAGCGCTCAATAAAGATCAATTCATAAATGTATCTTTCAATTTTAGGATTAACATAGT 300
Db      |||
```

Db 241 AACGCCCTCAATAAAAGATCAATTCATTAATGTAATCTTTCAATTTTAGGATAACAATACGT 300
Qy 301 GAACAGGGTATATTTTAAACGGTCAACAAATCTTAATAATTTTACCTGGCCGGTGAACA 360
Db 301 GAACAGGGTATATTTTAAACGGTCAACAAATCTTAATAATTTTACCTGGCCGGTGAACA 360
Qy 361 CCGTCTTCCAAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Db 361 CCGTCTTCCAAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Qy 421 AGGACGCTTAGGTGAATACACATTTGACTGTAGTCTTTTAAACAAAGAAACAAGTGGTTC 480
Db 421 AGGACGCTTAGGTGAATACACATTTGACTGTAGTCTTTTAAACAAAGAAACAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTATCTCAGCTACTATCTTT 540
Db 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTATCTCAGCTACTATCTTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTTAAATTAATGCCCATCCAAGGATAAGT 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTTAAATTAATGCCCATCCAAGGATAAGT 600
Qy 601 AAAATTCGGTAAACCGATTTGTTAATATATATATGTTTACACTTACAAGGATATTCGT 660
Db 601 AAAATTCGGTAAACCGATTTGTTAATATATATATGTTTACACTTACAAGGATATTCGT 660
Qy 661 AATACTTTTACGACAGAGACTTAGGTCAAAATGGAGCTGCTGTAACAGCCCTAGACT 720
Db 661 AATACTTTTACGACAGAGACTTAGGTCAAAATGGAGCTGCTGTAACAGCCCTAGACT 720
Qy 721 TGGTCACTGATAAATAGATAATTTGTTAGTAAATATATAGTAGGATCTACAATGACATTTAA 780
Db 721 TGGTCACTGATAAATAGATAATTTGTTAGTAAATATATAGTAGGATCTACAATGACATTTAA 780
Qy 781 ATTAGAGCTAATTAATAGATTAATAATAAAGAGAGGTAGTAAACAGAAAGCAGGTA 840
Db 781 ATTAGAGCTAATTAATAGATTAATAATAAAGAGAGGTAGTAAACAGAAAGCAGGTA 840
Qy 841 AAAACAAGAGCTCTGCTGTGTTAGTTAGTTGTTGAGCTCAATTTTAAAGTAATG 900
Db 841 AAAACAAGAGCTCTGCTGTGTTAGTTAGTTGTTGAGCTCAATTTTAAAGTAATG 900
Qy 901 TAAACTGATCTAAAGCAGATAGAAATTTAGTACAGGTTAAACCTTTTCAAGAAATTTATA 960
Db 901 TAAACTGATCTAAAGCAGATAGAAATTTAGTACAGGTTAAACCTTTTCAAGAAATTTATA 960
Qy 961 TTAACGAAATCAATTTTAAACATGCTCTCGCTGCTCAATATAAATAGGATCACTTAC 1020
Db 961 TTAACGAAATCAATTTTAAACATGCTCTCGCTGCTCAATATAAATAGGATCACTTAC 1020
Qy 1021 TGATCATCCATTAACCTGTTAAACAAATTCATGAGATAAATATCTTACAATGAA 1080
Db 1021 TGATCATCCATTAACCTGTTAAACAAATTCATGAGATAAATATCTTACAATGAA 1080
Qy 1081 AAGAAGGACATGCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCCCTCTGAAATGT 1140
Db 1081 AAGAAGGACATGCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCCCTCTGAAATGT 1140
Qy 1141 ATACATATGGATTCGACACGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
Db 1141 ATACATATGGATTCGACACGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
Qy 1201 AAAGAGAAAGAAAGTGGTAAAGTACGGGACCCCAATATATAATTTGATAGATTTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGTAAAGTACGGGACCCCAATATATAATTTGATAGATTTAG 1260
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTATATATATAAAAAATTTACTATTTG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTATATATAAAAAATTTACTATTTG 1320
Qy 1321 AGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
Db 1321 AGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
Qy 1376 AGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1376

Qy 1381 TAAATGGCAGAGGGAGTAATACCTTTATGATATATAAAATTTTGTATTTTGTATTCA 1440
Db 1377 TAAATGGCAGAGGGAGTAATACCTTTATGATATATAAAATTTTGTATTTTGTATTCA 1436
Qy 1441 TAAGATTATAAATCTATGTTTATAATGATAATAAATTTTAAAAATAAATACTATATAAT 1500
Db 1437 TAAGATTATAAATCTATGTTTATAATGATAATAAATTTTAAAAATAAATACTATATAAT 1496
Qy 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATCTAGAGTAATATGAAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATCTAGAGTAATATGAAATAAATCAG 1556
Qy 1561 TTATCTGAAAGCAATAATATCTTTTAAACAGCGTTCGGTCAAAATGGGAAGTTCATG 1620
Db 1557 TTATCTGAAAGCAATAATATCTTTTAAACAGCGTTCGGTCAAAATGGGAAGTTCATG 1616
Qy 1621 TGTAATCAATAGTATTAATAAAGTAATTTTAAATTAATTTTGTATTGTTTTCAGA 1680
Db 1617 TGTAATCAATAGTATTAATAAAGTAATTTTAAATTAATTTTGTATTGTTTTCAGA 1676
Qy 1681 AATTTAAATTAATTTTATGAGCATGGGAAGTTTCAAGGATCATTTGAGCAGCAGTACT 1740
Db 1677 AATTTAAATTAATTTTATGAGCATGGGAAGTTTCAAGGATCATTTGAGCAGCAGTACT 1736
Qy 1741 GTTTGAACAATGTATGTCGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTGAACAATGTATGTCGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATTTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACG 1860
Db 1797 GCATTTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACG 1856
Qy 1861 GATTTGATCTTTTCTTAACTTTTAAATTAATAAATAAATGATTTATGTAATTTATC 1920
Db 1857 GATTTGATCTTTTCTTAACTTTTAAATTAATAAATAAATGATTTATGTAATTTATC 1916
Qy 1921 AACACCTCAACATTTAGTGTAGGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
Db 1917 AACACCTCAACATTTAGTGTAGGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976
Qy 1981 CACATCAATCTTACACCAAACTTGAGCTTAATTTTCTACTTATCTCAGCAATAAC 2040
Db 1977 CACATCAATCTTACACCAAACTTGAGCTTAATTTTCTACTTATCTCAGCAATAAC 2036
Qy 2041 ATTCTAAATATC 2052
Db 2037 ATTCTAAATATC 2048

RESULT 6

AR076817 LOCUS AR076817 2042 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959176.

AR076817 ACCESSION AR076817

AR076817.1 GI:10003563

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2042)

AUTHORS Torikai,S. and Oeda,K.

TITLE Plant promoter and utilization thereof

JOURNAL Patent: US 5959176-A 2 28-SEP-1999;

FEATURES Location/Qualifiers

source

1..2042

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 11.8%; Score 241.2; DB 6; Length 2042;

Best Local Similarity 72.8%; Pred. No. 3.7e-28;

Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;

QY 1554 AATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAAAACAGCG-----TTCGGTCAAATG 1609
Db 1543 AATATTTATTTATCTGAATGATAACATCTTTGTAAACAAAACCTGGGCCAAATAGGACCATAA 1602
QY 1610 GGAAGTTTCATGTGATTTCAATAGTTTAAATATAAAAGTAAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCAAGTTTCAGTGATTTCTAAATGTTAAATACATGATGATTTTCCT--TTCAAGGT 1660
QY 1670 TTTGTTTCAGAAATTTAAAAATAAATATTGAGCATGGGAAAGTTCAAGGGCATCAATTGAGC 1729
Db 1661 ATAAGTTAAATCTTCAATCAATTAACCTTTAAATTTGGACATTAATTGAGCAACTTTATGCC 1720
QY 1730 AGCACTAGACTGTTTGACCAATGATGTCGGGTGTACATCTATGACCTTTCAACTCAAC 1789
Db 1721 CAGGTTGATTTGTTAAACACGTTTGTCCGGTGATATTAATTGACCTTTCAACTCAAGC 1780
QY 1790 TAGTGAAT-AATGATTTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGNATATATCTTTGAAATTTCAACAAACACAGCACTAACT 1840
QY 1849 TTTCTTTCAACGGATGGAATCCTTTTCTAAGATCTTTTAAATTTTAAATTAATGCAATTAAT 1908
Db 1841 TTTCTTTTAAACAGATTAGAAATCGTTTCTTAACTTTTAAATTT--AAAAAATAACATTAAT 1898
QY 1909 GTAATATTATCAACACTCAACATTCATGTTAGCTAGTACTATAAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTATCAACACTCAACATTCATGTTAGCTAGTACTATAAATAGGTGCTCTTGGTG 1958
QY 1969 CTCTACTATCATCATCAATCAATCTTACACCAACACCTTTGAGCTTAATTTTCTACTTAAT 2028
Db 1959 CTCTACTATCATCATCAATCTTCCAGCAACACCTTTGAGCTTAATTTTCTACTTAAT 2018
QY 2029 CTCAGCAATAACATCTTAAATATC 2052
Db 2019 TTTAGCAAAAACATCTTAAAGGTC 2042

RESULT 7
E15125
LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/2
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2042 /organism='Daucus carota L.'
FT promoter 1<..2042.
FT Location/Qualifiers
1..2042
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 11.8%; Score 241.2; DB 6; Length 2042;
Best Local Similarity 72.8%; Pred. No. 3.7e-28; Indels 9; Gaps 4;
Matches 367; Conservative 0; Mismatches 128;
QY 1554 AATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAAAACAGCG-----TTCGGTCAAATG 1609
Db 1543 AATATTTATTTATCTGAATGATAACATCTTTGTAAACAAAACCTGGGCCAAATAGGACCATAA 1602
QY 1610 GGAAGTTTCATGTGATTTCAATAGTTTAAATATAAAAGTAAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCAAGTTTCAGTGATTTCTAAATGTTAAATACATGATGATTTTCCT--TTCAAGGT 1660
QY 1670 TTTGTTTCAGAAATTTAAAAATAAATATTGAGCATGGGAAAGTTCAAGGGCATCAATTGAGC 1729
Db 1661 ATAAGTTAAATCTTCAATCAATTAACCTTTAAATTTGGACATTAATTGAGCAACTTTATGCC 1720
QY 1730 AGCACTAGACTGTTTGACCAATGATGTCGGGTGTACATCTATGACCTTTCAACTCAAC 1789
Db 1721 CAGGTTGATTTGTTAAACACGTTTGTCCGGTGATATTAATTGACCTTTCAACTCAAGC 1780
QY 1790 TAGTGAAT-AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGNATATATCTTTGAAATTTCAACAAACACAGCACTAACT 1840
QY 1849 TTTCTTTCAACGGATGGAATCCTTTTCTAAGATCTTTTAAATTTTAAATTAATGCAATTAAT 1908
Db 1841 TTTCTTTTAAACAGATTAGAAATCGTTTCTTAACTTTTAAATTT--AAAAAATAACATTAAT 1898
QY 1909 GTAATATTATCAACACTCAACATTCATGTTAGCTAGTACTATAAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTATCAACACTCAACATTCATGTTAGCTAGTACTATAAATAGGTGCTCTTGGTG 1958
QY 1969 CTCTACTATCATCATCAATCTTACACCAACACCTTTGAGCTTAATTTTCTACTTAAT 2028
Db 1959 CTCTACTATCATCATCAATCTTCCAGCAACACCTTTGAGCTTAATTTTCTACTTAAT 2018
QY 2029 CTCAGCAATAACATCTTAAATATC 2052
Db 2019 TTTAGCAAAAACATCTTAAAGGTC 2042

RESULT 8

AR076816
LOCUS AR076816 247 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 1 from patent US 5959176.
ACCESSION AR076816
VERSION AR076816.1 GI:10003562
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 1 28-SEP-1999;
FEATURES
1..247
source /organism='unknown'
/mol_type='unassigned DNA'

ORIGIN

Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.9e-21; Indels 2; Gaps 1;
Matches 223; Conservative 0; Mismatches 24;
QY 1804 TTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1863
Db 1 TTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACAGAT 60
QY 1864 TGAATCTTTTCTTCAAACTTTTAAATAAATAAATAATGCAATTAATGTAATTAATTAATCAAC 1923

Db 61 TAGAATCGTTCTTCTAAACCTTTTAAATTT--AAAAAATACATTACTATAATATTTATCAAC 118
Qy 1924 ACCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1983
Db 119 ACCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178
Qy 1984 ATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAATAACATT 2043
Db 179 ATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTAAATTTTAGCAAAAACATT 238
Qy 2044 CTTAAATATC 2052
Db 239 CTTAAAGGTC 247

RESULT 9
E15124
LOCUS E15124 247 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15124
VERSION E15124.1 GI:5709807
KEYWORDS JP 1998052273-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 1 24-FEB-1998;
COMMENT SUMITOMO CHEM CO LTD
OS Daucus carota L. (carrot)
PN JP 1998052273-A/1
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH
FH
FT source 1..247
FT
FT
FT promoter 1<..<247.
FT Location/Qualifiers
FT 1..247
FT /organism="unidentified"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:32644"

ORIGIN
Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.9e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
Qy 1804 TTCTAGAATACATCTTTTCAAAATTTCAACAAACAGCGTTTAACTTTTCTTTCAACGGAT 1863
Db 1 TTCTAGAATATATCTTTTGAATTTTCAACAAACAGCGCTAACTTTTCTTTTAAACAGAT 60
Qy 1864 TGGAAATCCTTTCTTAAACCTTTTAAATATAAAAAATGCAATATTGTAATATTTATCAAC 1923
Db 61 TAGAATCGTTCTTAAACCTTTTAAATTT--AAAAAATACATTACTATAATATTTATCAAC 118
Qy 1924 ACCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1983
Db 119 ACCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178
Qy 1984 ATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAATAACATT 2043
Db 179 ATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTAAATTTTAGCAAAAACATT 238

Qy 2044 CTTAAATATC 2052
Db 239 CTTAAAGGTC 247

RESULT 10
E15124
LOCUS E15124 246 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 2 from patent US 6218598.
ACCESSION E15124
VERSION E15124.1 GI:15110041
KEYWORDS .
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,F., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: US 6218598-A 2 17-APR-2001;
FEATURES Location/Qualifiers
source 1..246
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
Qy 1805 TCTAGAATACATCTTTTCAAAATTTCAACAAACAGCGTTTAACTTTTCTTTCAACGGATT 1864
Db 1 TCTAGAATATATCTTTTGAATTTTCAACAAACAGCGCTAACTTTTCTTTTAAACAGATT 60
Qy 1865 GGAATCCTTTTCTAAACCTTTTAAATATAAAAAATGCAATATTGTAATATTTATCAACA 1924
Db 61 AGAATCGTTTCTTAAACCTTTTAAATTT--AAAAAATACATTACTATAATATTTATCAACA 118
Qy 1925 CCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1984
Db 119 CCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178
Qy 1985 TCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAATAACATT 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATTTTCTACTAAATTTTAGCAAAAACATT 238
Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 11
E55065
LOCUS E55065 246 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter.
ACCESSION E55065
VERSION E55065.1 GI:18625251
KEYWORDS JP 2000083679-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,I., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: JP 2000083679-A 2 28-MAR-2000;
COMMENT SUMITOMO CHEM CO LTD
OS Daucus carota L.
PN JP 2000083679-A/2
PD 28-MAR-2000
PF 12-JUL-1999 JP 1999197240
PR IKUHARU ISHIGE,SATOMI NISHIKAWA,KENJI OEDA
PI C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N15/09,C12R1:91), PC
PC (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91), PC

(C12N5/00, C12R1:91)	
CC	Location/Qualifiers
FT	1. .246
FT	/organism="Daucus carota L."
FEATURES	
source	Location/Qualifiers
	1. .246
	/organism="unidentified"
	/mol_type="genomic DNA"
	/db_xref="taxon:32644"
ORIGIN	
Query Match 9.5%; Score 195.6; DB 6; Length 246;	
Best Local Similarity 89.5%; Pred. No. 7e-21; Mismatches 24; Indels 2; Gaps 1;	
Matches 222; Conservative 0;	
Qy	1805 TCTAGATACATCTTTTCAAAATTTCAACAAACAGAGCTTTAACTTTTCTTTCAACGATT 1864
Db	1 TCTAGATATATCTTTTGAATTTTCAACAAACAGAGCTTTAACTTTTCTTTTAAACAGATT 60
Qy	1865 GGAATCTTTTCTAACTTTTAAATAAAAAAATGCATTTATGTAATTTATCAACA 1924
Db	61 AGATCGTTTTCTAACTTTTAAATTT--AAAAATACATTTACTATATATTTATCAACA 118
Qy	1925 CCTCAACATTGATCTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCA 1984
Db	119 CCTCAACATTGATCTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCA 178
Qy	1985 TCAATCTTACACCAACAACTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATAACATTC 2044
Db	179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAATTTTGTAGCAAAACATTC 238
Qy	2045 TAAATATC 2052
Db	239 TAAAGGTC 246
RESULT 12	
DCAL8706/c	8005 bp DNA linear PLN 10-FEB-1999
LOCUS	DCAL8706
DEFINITION	Daucus carota Inv*Dc5 gene.
ACCESSION	Y18706
VERSION	Y18706.1 GI:4454114
KEYWORDS	beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform.
SOURCE	Daucus carota (carrot)
ORGANISM	Daucus carota
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae; Daucus.
REFERENCE	1
AUTHORS	Sturm, A.
TITLE	Molecular characterisation and functional analysis of sucrose-cleaving enzymes in carrot (Daucus carota L.)
JOURNAL	J. Exp. Bot. 47, 1187-1192 (1996)
REFERENCE	2 (bases 1 to 8005)
AUTHORS	Sturm, A.
TITLE	Direct Submission
JOURNAL	Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut, Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND
COMMENT	Related sequence X67163.
FEATURES	Location/Qualifiers
source	1. .8005
	/organism="Daucus carota"
	/mol_type="genomic DNA"
	/cultivar="Nantaise"
	/db_xref="taxon:4039"
	1. .8005
	/genes="Inv*Dc5"
gene	1. .2666
5'UTR	join(2767..3160,3724..3732,4656..5515,5605..5766,5869..6104,6190..6277,6384..6587)
CDS	
/gene="Inv*Dc5"	
/EC_number="3.2.1.26"	
/note="soluble acid invertase"	
/codon_start=1	
/product="beta-fructofuranosidase, isoform II"	
/protein_id="CAA77266.1"	
/db_xref="GI:4454115"	
/db_xref="GOA:Q42722"	
/db_xref="UniProt/TREMBL:Q42722"	
/translation="MEHPITISHYTPDPGEHSPSLTTNTAQSSRRRLTFFVLLFS	
SILAACLVMGTWLPFNSGNEAVEKSTVPEETVEVAPRGVAGVSGMKSFRPALNAE	
PPANFPWNSVLWSQSSSFHQPNQNMNDPNGLFYKGYHLFYQVNPDAWGNKI	
VHGAVSSDLIHWKHLPVAMVTDHWYDNGVWTSATILPDGOIWMLYTGSTNESVOV	
QNLAYPADSPDLLIEWKYPGNPVLVPPPGIDFKDFRDETTRWTFEGKRWLIIGSK	
LNKTGISLVDTVDVFNFTLLDGLVHVGHTGMECVDFYFVSKFGENGIDTISFDVG	
VGHVMAKSLDDDRNDYYAIGTYPVSGKWVPDNPDLVDVIGLGYDYGIYASKTFFYS	
NKKRRVLWNIKETDSEISDVRKGMASVOQIPRTILFDPKTGSNLLQWPVEVKNKRL	
NKTVPFNVEINTGAVLPLBIGSGQLDTAEFVVDKESLERVOETRESVYDCKNGGSS	
GRGALGPFGLILLADKDLSEOTPVFYIAGKSGGNLRTFFCADHSRSSKAVDVKETV	
GSVVPVLRGEKLTMRILVDHSIVESFSQGRCTISRVYPTKAIYNNAKVFLFNNAE	
ARIASLINWQMTAQRTQTHFADLVI"	
2767..3160	
/genes="Inv*Dc5"	
/number=1	
3161..3723	
/genes="Inv*Dc5"	
/number=1	
3724..3732	
/genes="Inv*Dc5"	
/number=2	
3733..4655	
/genes="Inv*Dc5"	
/number=2	
4656..5515	
/genes="Inv*Dc5"	
/number=3	
5516..5604	
/genes="Inv*Dc5"	
/number=3	
5605..5766	
/genes="Inv*Dc5"	
/number=4	
5767..5868	
/genes="Inv*Dc5"	
/number=4	
5869..6104	
/genes="Inv*Dc5"	
/number=5	
6105..6189	
/genes="Inv*Dc5"	
/number=5	
6190..6277	
/genes="Inv*Dc5"	
/number=6	
6278..6383	
/genes="Inv*Dc5"	
/number=6	
6384..6587	
/genes="Inv*Dc5"	
/number=7	
6588..8005	
/genes="Inv*Dc5"	
3'UTR	
ORIGIN	
Query Match 8.5%; Score 174; DB 8; Length 8005;	
Best Local Similarity 81.7%; Pred. No. 8.2e-18;	
Matches 201; Conservative 0; Mismatches 45; Indels 0; Gaps 0;	
Qy	17 CACATAGGGCCCTGTTGGTTGAGAGAGAGAGCTTCTGACTTCTTCTTTTGA 76
Db	1285 CACTATGGGGGTGTTTGTTCCTCCGGAAGTAGGAGCTGCTTCGCTTCTCTTGA 1236
Qy	77 CCTGTTTGTATAAGAAGTAGAATAATTTTAAAGCTCGGAATACTTCTCTCTC 136

Db 1225 CCCGTTTGTGTAATAAGCAGACATTTTAAAAAGTTGAGATGCTAGCTTCTCTC 1166
 Qy 137 ACAACTCCGCTCTTTTCCAAACACTTTATTAACCTTTTACTTCTCACTTCTCA 196
 Db 1165 ACAGCTTCGCTCTTTTCCAAACACTTTATTAATTTATTTACTTCTCACTTCT 1106
 Qy 197 CTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAA 256
 Db 1105 CTTTCTTAATTTAAGTAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTATGACT 1046
 Qy 257 GATCAT 262
 Db 1045 GACCCCT 1040

RESULT 13
 BD188672
 LOCUS BD188672 2831 bp DNA linear PAT 17-JUL-2003
 DEFINITION Promotor and use thereof.
 ACCESSION BD188672
 VERSION BD188672.1 GI:32998411
 KEYWORDS JP 2003000252-A/1.
 SOURCE unidentified
 ORGANISM unclassified.

REFERENCE 1 (bases 1 to 2831)
 AUTHORS Nishikawa,S. and Oeda,K.
 TITLE Promotor and use thereof
 JOURNAL Patent: JP 2003000252-A 1 07-JAN-2003;
 SUMITOMO CHEMICAL CO LTD
 COMMENT OS Daucus carota L. (carrot)
 PN JP 2003000252-A/1
 PD 07-JAN-2003
 PF 31-MAY-2001 JP 2001164069
 PI SATOMI NISHIKAWA,KENJI OEDA
 PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/00,C12N5/00
 CC Promotor and use thereof
 FH Key Location/Qualifiers
 FT promoter (1). (2831).

FEATURES
 source
 1. .2831
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 7.9%; Score 162.4; DB 6; Length 2831;
 Best Local Similarity 82.5%; Pred. No. 6.4e-16;
 Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
 Qy 22 AGGCCCTGTTGGTTGAGAGAGAGAGCTGCTTCTGACTTCTCTCTTTTGACCTGT 81
 Db 1716 AGGCCCTGTTGGTTTATGGAATCAGAGCTGCTTCTGACTTCTGCTTTT-ACCGGT 1774
 Qy 82 TTGTATAAAGAGTAGAAATATTTTAAAAAGCTGCGAATCACTTCTCTCTCACAA 141
 Db 1775 TTGTATAAAGAGAGAGACTTTTAAAGAGCTGAGAGTCTAGTTCTCTCACAG 1834
 Qy 142 TTCGGCTCTTTTCCAAACACTTTATTAACCTTTTACTTCTCAATTTCTACTCCACTTCT 201
 Db 1835 TTCTGCTCTTTTCCAAACACTTTATCACTTACTTCTCACTTCTGCTTCTCTCTCT 1894
 Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAGATCA 261
 Db 1895 TTACTATAAGCAAGAGTCAATTTCTTTTAAATTAACCCAAACGGCCTTAAGTAATTGA 1954

RESULT 14
 BD188674
 LOCUS BD188674 2865 bp DNA linear PAT 17-JUL-2003
 DEFINITION Promotor and use thereof.
 ACCESSION BD188674
 VERSION BD188674.1 GI:32998413
 KEYWORDS JP 2003000252-A/3.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 2865)
 AUTHORS Nishikawa,S. and Oeda,K.
 TITLE Promotor and use thereof
 JOURNAL Patent: JP 2003000252-A 3 07-JAN-2003;
 SUMITOMO CHEMICAL CO LTD
 COMMENT OS Daucus carota L. (carrot)
 PN JP 2003000252-A/3
 PD 07-JAN-2003
 PF 31-MAY-2001 JP 2001164069
 PI SATOMI NISHIKAWA,KENJI OEDA
 PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/00,C12N5/00
 CC Promotor and use thereof
 FH Key Location/Qualifiers
 FT promoter (1). (2865).

FEATURES
 source
 1. .2865
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 7.9%; Score 162.4; DB 6; Length 2831;
 Best Local Similarity 82.5%; Pred. No. 6.4e-16;
 Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
 Qy 22 AGGCCCTGTTGGTTGAGAGAGAGAGCTGCTTCTGACTTCTCTCTTTTGACCTGT 81
 Db 1716 AGGCCCTGTTGGTTTATGGAATCAGAGCTGCTTCTGACTTCTGCTTTT-ACCGGT 1774
 Qy 82 TTGTATAAAGAGTAGAAATATTTTAAAAAGCTGCGAATCACTTCTCTCTCACAA 141
 Db 1775 TTGTATAAAGAGAGAGACTTTTAAAGAGCTGAGAGTCTAGTTCTCTCTCACAG 1834
 Qy 142 TTCGGCTCTTTTCCAAACACTTTATTAACCTTTTACTTCTCAATTTCTACTCCACTTCT 201
 Db 1835 TTCTGCTCTTTTCCAAACACTTTATCACTTACTTCTCACTTCTGCTTCTCTCTCT 1894
 Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAGATCA 261
 Db 1895 TTACTATAAGCAAGAGTCAATTTCTTTTAAATTAACCCAAACGGCCTTAAGTAATTGA 1954

ACCESSION
 VERSION BD188674.1 GI:32998413
 KEYWORDS JP 2003000252-A/3.
 SOURCE unidentified
 ORGANISM unclassified.

REFERENCE 1 (bases 1 to 2865)
 AUTHORS Nishikawa,S. and Oeda,K.
 TITLE Promotor and use thereof
 JOURNAL Patent: JP 2003000252-A 3 07-JAN-2003;
 SUMITOMO CHEMICAL CO LTD
 COMMENT OS Daucus carota L. (carrot)
 PN JP 2003000252-A/3
 PD 07-JAN-2003
 PF 31-MAY-2001 JP 2001164069
 PI SATOMI NISHIKAWA,KENJI OEDA
 PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/00,C12N5/00
 CC Promotor and use thereof
 FH Key Location/Qualifiers
 FT promoter (1). (2865).

FEATURES
 source
 1. .2865
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN

Query Match 7.9%; Score 162.4; DB 6; Length 2865;
 Best Local Similarity 82.5%; Pred. No. 6.4e-16;
 Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
 Qy 22 AGGCCCTGTTGGTTGAGAGAGAGAGCTGCTTCTGACTTCTCTCTTTTGACCTGT 81
 Db 1716 AGGCCCTGTTGGTTTATGGAATCAGAGCTGCTTCTGACTTCTGCTTTT-ACCGGT 1774

Qy 82 TTGTATAAAGAGTAGAAATATTTTAAAAAGCTGCGAATCACTTCTCTCTCACAA 141
 Db 1775 TTGTATAAAGAGAGAGACTTTTAAAGAGCTGAGAGTCTAGTTCTCTCTCACAG 1834
 Qy 142 TTCGGCTCTTTTCCAAACACTTTATTAACCTTTTACTTCTCAATTTCTACTCCACTTCT 201
 Db 1835 TTCTGCTCTTTTCCAAACACTTTATCACTTACTTCTCACTTCTGCTTCTCTCTCT 1894

Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAGATCA 261
 Db 1895 TTACTATAAGCAAGAGTCAATTTCTTTTAAATTAACCCAAACGGCCTTAAGTAATTGA 1954

RESULT 15
 DARGCHS2
 LOCUS DARGCHS2 4886 bp DNA linear PLN 14-APR-2000
 DEFINITION Carrot gCHS2 gene for chalcone synthase.
 ACCESSION D16255
 VERSION D16255.1 GI:441168
 KEYWORDS chalcone synthase.
 SOURCE Daucus carota (carrot)
 ORGANISM Daucus carota

Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Apiales; Apiaceae; Scandiceae;
 Daucinae; Daucus.

REFERENCE 1 (bases 1 to 4886)
 AUTHORS Ozeki,Y., Davies,E. and Takeda,J.
 TITLE Structure and expression of chalcone synthase gene in carrot
 suspension cultured cells regulated by 2,4-D
 JOURNAL Plant Cell Physiol. 34, 1029-1037 (1993)
 REFERENCE 2 (bases 1 to 4886)
 AUTHORS Ozeki,Y
 TITLE Direct Submission

Submitted (14-MAY-1993) Yoshihiro Ozeki, College of Arts and
 Sciences, The University of Tokyo, Department of Biology; Komaba,
 Meguro-ku, Tokyo 153, Japan (Tel:03-3467-1171(ex.253),

Search completed: October 27, 2005, 21:12:07
Job time : 9014 secs

COMMENT Submitted (14-MAY-1993) to DDBJ by:
Yoshihiro Ozeki
Department of Biol. College of
Arts & Science, Univ. of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153
Japan
Phone:
Fax:

Fax: 03-3485-2904)

03-3467-1171 x253

03-3485-2904.

Location/Qualifiers

1..4886
/organism="Daucus carota"
/mol_type="genomic DNA"
/cultivar="Kurodagosun"
/db_xref="taxon:4039"
/dev_stage="cultured cells"
1957..1962
/note="TACPyAT motif"
2060..2065
/note="G-box"
2286..2291
/note="G-box"
2303..2308
/note="TACPyAT motif"
2327..2331
TATA_signal
prim_transcript 2358..4519
/note="CHS2 mRNA and intron"
2358..2607
/number=1
join(2430..2607,3395..4386)
/gene="gCHS2"
join(2430..2607,3395..4386)
/gene="gCHS2"
/EC_number="2.3.1.74"
/codon_start=1
/product="chalcone synthase"
/protein_id="BAA03784.1"
/db_xref="GI:565639"
/translation="MVTVNFRAQRAEGPATVLAIGTATPPNCVDQSAAYVYFRIT
NSEDKPELKKERMCCKSMINTRYMHLTEDLLKONPSFCEYMASSLDARQDIVNNEV
PKLGEALRALKEWGPQSKLTHLJFCTTSGVDMPCGADRLTKLLGLRPSVKRFMY
QQCGFAGTVLRLAKDLAENKNARLVVCSSEIVITFRGPNTHLDSLVGQALFGDG
AGAVVGSDFPVGIEKPLFEIVSAAQTILPDSGAIDGHLREVLGTLPHLLKDVPLGIS
KNIRKSLVEAFKPIGIDWNSIFWIAHPPGPAILDQVETLSLKEPKLSTQVLRDY
GNMSSACVLFILDEMRSKASANDGHRITTEGGLDWGLFPGFGLTVETVVLHVSPT"

intron

2608..3394
/gene="gCHS2"

3395..4519

/number=2

polyA_signal 4501..4506

ORIGIN

Query Match 7.9%; Score 162.4; DB 8; Length 4886;
Best Local Similarity 82.5%; Pred. No. 5.8e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
Qy 11 CTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAAGCTGCTTCTGACTTCTTCTTC 70
Db 1355 CTAGTAGATGTAGGCGCTGTTGGATGGTGAAGCAGAGCTGCTTCGGGCTTCTGCTTT 1414
Qy 71 TTTTGACCTGTTGTATAGAAAGTAGAAATATTTTAAAGCTGGGAATACATACTTC 130
Db 1415 TCTTGACCGTTTGGGTAAGAGTAGAAGCACTTTTAAAGAGTTGAGAATACATGCTTC 1474
Qy 131 TCTCTCAACTCTCG-CTTCTTTTCCAAACACTTTTAACTTTTACTTCTCATTTTC 189
Db 1475 TCTCTCAGAGCTTCTGTTTTTTTTTTCNAACACTTTATTTATTTTACTTTTCACTTC 1534
Qy 190 TACTCCACTTCTTCTATAAGCAAGAAATACATTTCTTTTAAAGCTAACCCAAACGGCTC 249
Db 1535 TACTCCACTTCTTTAAATTTAAGCAAGAGTACATTTCTTTTAACTAACCCAAACGGCCCC 1594

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2005, 15:39:45 ; Search time 1123 Seconds
(without alignments)
10816.838 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtgcctacagcaca.....gcaataacattctaataatc 2052

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2052	100.0	2052	3 AAA37959	Aaa37959 Carrot pr
2	2048.8	99.8	2052	3 AAA37961	Aaa37961 Carrot pr
3	2034.8	99.2	2056	3 AAA37964	Aaa37964 Plasmid #
4	2030.8	99.0	2048	3 AAA37962	Aaa37962 Plasmid #
5	2030.8	99.0	2048	3 AAA37963	Aaa37963 Plasmid #
6	239.6	11.7	2042	2 AAV15144	Aav15144 New promo
7	196.6	9.6	247	2 AAV15143	Aav15143 New promo
8	195.6	9.5	246	3 AAZ49611	Aaz49611 Carrot CR
9	194.4	9.5	196	12 ADP07499	Adp07499 Carrot DN
10	182.4	7.9	2831	10 ADC56759	Adc56759 Carrot DN
11	162.4	7.9	2865	10 ADC56761	Adc56761 Carrot DN
12	119	5.8	140	3 AAZ49616	Aaz49616 Oligonucle
13	119	5.8	140	3 AAZ49615	Aaz49615 Oligonucle
14	99.4	4.8	8056	8 ABZ10246	Abz10246 Haematopo
15	91.8	4.5	8056	8 ABZ10100	Abz10100 Haematopo
16	91.6	4.5	8056	8 ABZ10246	Abz10246 Haematopo
17	88.6	4.3	851	3 AAA37960	Aaa37960 Carrot te
18	88.6	4.3	8056	8 ABZ10100	Abz10100 Haematopo
19	87.2	4.2	110000	13 ABD32968	Continuation (7 of
20	83	4.0	158001	12 ADL17884	Adl17884 Human pho

C 21	75	3.7	11745	6	ABK28332	Abk28332 DNA trans
C 22	71.2	3.5	8222	8	ACF62816	Acf62816 Colon can
C 23	71.2	3.5	11222	10	ADB54318	Adb54318 Pretreat
C 24	71	3.5	5930	6	ABL32517	Ab132517 Human imm
C 25	70.8	3.5	5286	13	ADS89278	Ads89278 Oligonucle
C 26	70.8	3.5	5286	13	ADS89552	Ads89552 Oligonucle
C 27	69.8	3.4	6352	6	ABK31340	Abk31340 Signal tr
C 28	69.8	3.4	6352	6	ABL70563	Ab170563 Chemical
C 29	69.8	3.4	6352	6	AAS61235	Aas61235 Human gen
C 30	69.6	3.4	8222	8	ACF62794	Acf62794 Colon can
C 31	69.6	3.4	11222	10	ADB54190	Adb54190 Pretreat
C 32	69.4	3.4	16258	6	ABK40038	Abk40038 Human che
C 33	69.4	3.4	16258	6	ABL70376	Ab170376 Chemical
C 34	69.2	3.4	6045	6	ABK31541	Abk31541 Signal tr
C 35	69.2	3.4	6045	6	ABL70624	Ab170624 Chemical
C 36	68.2	3.3	15548	6	ABL34155	Ab134155 Human imm
C 37	68.2	3.3	110000	13	ABD32968	Continuation (7 of
C 38	68	3.3	11422	6	ABK39936	Abk39936 Human che
C 39	68	3.3	11422	6	ABL32218	Ab132218 Human imm
C 40	67.8	3.3	15732	4	AAS45388	Aas45388 Chemical
C 41	67.8	3.3	15732	6	ABK28233	Abk28233 DNA trans
C 42	67.8	3.3	99784	13	ABD32992	Abd32992 Human can
C 43	67.6	3.3	5984	6	ABQ66994	Abq66994 Human ang
C 44	67.6	3.3	6050	6	ABL34011	Ab134011 Human imm
C 45	67.4	3.3	5286	13	ADS89551	Ads89551 Oligonucle

ALIGNMENTS

RESULT 1

AAA37959
ID AAA37959 standard; DNA; 2052 BP.

AC AAA37959;

DT 18-AUG-2000 (first entry)

DE Carrot promoter sequence #1.

KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

OS Daucus carota.

PN W0200020613-A1.

PD 13-APR-2000.

PF 28-SEP-1999; 99WO-JP005303.

PR 02-OCT-1998; 98JP-00281124.

PA (SUMO) SUMITOMO CHEM CO LTD.

PI Nishikawa S, Oeda K;

DR WPI; 2000-303791/26.

PT New Plant promoters and terminators from Daucus carota L., useful in

plant breeding, for e.g. controlling fertilities of plants.

PS Claim 1; Page 69-70; 81pp; English.

CC This sequence represents a carrot promoter. The invention relates to plant promoters and terminators from Daucus carota L. which are capable of expressing a gene of interest in plants. The invention also includes a chimeric gene characterized in that it comprises the promoter and a desired gene linked to each other in the form capable of functioning. A method of producing a transformant comprises introducing the promoter, the chimeric gene or a vector comprising the promoter and a desired gene or terminator sequence into a host cell. The plant promoters and terminators are useful in plant breeding, for e.g. fertilities of plants may be controlled by expressing, in the host cells, a sense or antisense

CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 2052; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 CATGTGCGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAGAAAGCTGCTCTGA 60
Db 1 CATGTGCGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAGAAAGCTGCTCTGA 60
Qy 61 CTTCCTCTCTTTTGACCTGTTGTATATAAGAACTGAGAAATATTTTAAAGCTGCGAA 120
Db 61 CTTCCTCTCTTTTGACCTGTTGTATAAGAAAGTAGAAATATTTTAAAGCTGCGAA 120
Qy 121 TACTAACTTCTCTCTCACAACTCCGCTTCTTTTCCAAACACTTTATTAACCTTTTACT 180
Db 121 TACTAACTTCTCTCTCACAACTCCGCTTCTTTTCCAAACACTTTATTAACCTTTTACT 180
Qy 181 TCTCATTTCTACTCCACTTCTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
Db 181 TCTCATTTCTACTCCACTTCTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
Qy 241 AACGGCTCAATAAAGATCATTCATAAATGTATCTTTTCAATTTTAGGATAACAATACGT 300
Db 241 AACGGCTCAATAAAGATCATTCATAAATGTATCTTTTCAATTTTAGGATAACAATACGT 300
Qy 301 GAACAGGGTATTTTAAAGCTGTCAACAAATCTTAATAATTTTACCTGGCCGGTGAACA 360
Db 301 GAACAGGGTATTTTAAAGCTGTCAACAAATCTTAATAATTTTACCTGGCCGGTGAACA 360
Qy 361 CCGTCTTCCAAGATAATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCATGC 420
Db 361 CCGTCTTCCAAGATAATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCATGC 420
Qy 421 AGGACGACTTAGGTGAATACACATTTGACTGTGAGTCTTTAAACAAAGAAACAGTGGTTC 480
Db 421 AGGACGACTTAGGTGAATACACATTTGACTGTGAGTCTTTAAACAAAGAAACAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGACAAAACCCGACACAACTCTATCCACGTACTATACCTT 540
Db 481 ATGCTCAGCCATCAAAATTTGACAAAACCCGACACAACTCTATCCACGTACTATACCTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATGTAATGTAATGTAATGTAATGTA 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATGTAATGTAATGTAATGTAATGTA 600
Qy 601 AAAATTCCTGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATAT 660
Db 601 AAAATTCCTGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATAT 660
Qy 661 AATACTTTTACAGCAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCCTAGACT 720
Db 661 AATACTTTTACAGCAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCCTAGACT 720
Qy 721 TGGTCACTGATAAATAGTAATTTGTTAGTATATATATATATATATATATATATATATAT 780
Db 721 TGGTCACTGATAAATAGTAATTTGTTAGTATATATATATATATATATATATATATATAT 780
Qy 781 ATTAGAGCTATTAAATAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Db 781 ATTAGAGCTATTAAATAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Qy 841 AAAACAAGAGCTTCTGCTGCTGTTTGTAGTTGTTGCTGAGTCACTTTTAAAGTAATG 900
Db 841 AAAACAAGAGCTTCTGCTGCTGTTTGTAGTTGTTGCTGAGTCACTTTTAAAGTAATG 900
Qy 901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTCAAGAAATTTATA 960
Db 901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTCAAGAAATTTATA 960
Qy 961 TTAAACGAAATCATTTTATAACATGTCTCTCGGCTGCTCATTTAATAGGATCACTTAC 1020
```

```
Db 961 TTAAACGAAATCATTTTATAACATGTCTCTCGGCTGCTCATTTAATAGGATCACTTAC 1020
Qy 1021 TGATCATCCATTAAACACCTTGTAAACAAATCAATGAGATAAAATATCTTACAATGAA 1080
Db 1021 TGATCATCCATTAAACACCTTGTAAACAAATCAATGAGATAAAATATCTTACAATGAA 1080
Qy 1081 AGAAGGACAATGCTCTTTTGAAGAAACAAATAGGTACTCCCTCCGCTCTGAAATGT 1140
Db 1081 AGAAGGACAATGCTCTCTTTTGAAGAAACAAATAGGTACTCCCTCCGCTCTGAAATGT 1140
Qy 1141 ATACATATGATTTGGACACGAGACTAAAGAAATGTATAAAGTAAATGTAGAGTAAAG 1200
Db 1141 ATACATATGATTTGGACACGAGACTAAAGAAATGTATAAAGTAAATGTAGAGTAAAG 1200
Qy 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATATATATATATATATTTG 1260
Db 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATATATATATATATTTG 1260
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATATATATATATATTTG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATATATATATATTTG 1320
Qy 1321 AGAAAGTTTTGAAATGTATAGAAATTCGAGTGGGACATCCATAAAGGAAAGTGTAGAAT 1380
Db 1321 AGAAAGTTTTGAAATGTATAGAAATTCGAGTGGGACATCCATAAAGGAAAGTGTAGAAT 1380
Qy 1381 TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTGATTTCA 1440
Db 1381 TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTGATTTCA 1440
Qy 1441 TAAAGATTATAAATCTATGTTTATAATGATAATATAAATTTTAAATAATATATATTAAT 1500
Db 1441 TAAAGATTATAAATCTATGTTTATAATGATAATATAAATTTTAAATAATATATATTAAT 1500
Qy 1501 CTGATTAGTCGATTACCGCTTTTAAATTTTCAATACATGAGTAAATGTAATAAATCAG 1560
Db 1501 CTGATTAGTCGATTACCGCTTTTAAATTTTCAATACATGAGTAAATGTAATAAATCAG 1560
Qy 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCATG 1620
Db 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCATG 1620
Qy 1621 TGATTTCAATAGTTTTTAATAAAGTAAATTTTAAATTTTAAATTTTGTATTTTGTTCAGA 1680
Db 1621 TGATTTCAATAGTTTTTAATAAAGTAAATTTTAAATTTTAAATTTTGTATTTTGTTCAGA 1680
Qy 1681 AATTTAAATATAATTTTGAAGCATGGGAAGTTCACGGGCATCATTTGAGCAGCAGTAGACT 1740
Db 1681 AATTTAAATATAATTTTGAAGCATGGGAAGTTCACGGGCATCATTTGAGCAGCAGTAGACT 1740
Qy 1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAT 1800
Db 1741 GTTTGAACAATGTATGTCCGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAT 1800
Qy 1801 GCATTTCTAGAAATACATCTTTTCAAAATTTTCAACAAAACAGCTTTAACTTTTCTTCAACG 1860
Db 1801 GCATTTCTAGAAATACATCTTTTCAAAATTTTCAACAAAACAGCTTTTAACTTTTCTTCAACG 1860
Qy 1861 GATTTGGAATCCTTTTCTTAAACTTTTAAATAAATAAATAAATGCAATTTTGTAAATTTTATC 1920
Db 1861 GATTTGGAATCCTTTTCTTAAACTTTTAAATAAATAAATAAATGCAATTTTGTAAATTTTATC 1920
Qy 1921 AACACCTCAACATTTAGTTAGGCTACTATAAATAGGTGCTCTTTGGTGCTCTACTATCAT 1980
Db 1921 AACACCTCAACATTTAGTTAGGCTACTATAAATAGGTGCTCTTTGGTGCTCTACTATCAT 1980
Qy 1981 CACATCAATCTTACACCAAAACCTTTGAGCTTTAAATTTTCTTCTTCTCAGCAATAAC 2040
Db 1981 CACATCAATCTTACACCAAAACCTTTGAGCTTTAAATTTTCTTCTTCTCAGCAATAAC 2040
Qy 2041 ATTCTAAATATC 2052
Db 2041 ATTCTAAATATC 2052
```


QY 1381 TAAATGGACAGAGGGAGTAATACCTTTATGATATATAAAATTTTGTATTTGATTTC 1440
 Db 1381 TAAATGGACAGAGGGAGTAATACCTTTATGATATATAAAATTTTGTATTTGATTTC 1440
 QY 1441 TAAGATTATAAATCTATGTTTATAATGATAATATATATTTTAAAAATAATACTATATTAATT 1500
 Db 1441 TAAGATTATAAATCTATGTTTATAATGATAATATATATTTTAAAAATAATACTATATTAATT 1500
 QY 1501 CTGATTAGTCGATTACCGCCCTTTTATAATTTTACAATCTGAGTAATATGAATAAATCAG 1560
 Db 1501 CTGATTAGTCGATTACCGCCCTTTTATAATTTTACAATCTGAGTAATATGAATAAATCAG 1560
 QY 1561 TTATCTGAAGCAAAATAATCTTTGTATAACAGCGTTTGGTCAAAATGGAAGTTTCATG 1620
 Db 1561 TTATCTGAAGCAAAATAATCTTTGTATAACAGCGTTTGGTCAAAATGGAAGTTTCATG 1620
 QY 1621 TGTATTCAAATAGTTTATAATATAAAGTAAATTTTAAAAATTAATCTTATTTTGTTCAGA 1680
 Db 1621 TGTATTCAAATAGTTTATAATATAAAGTAAATTTTAAAAATTAATCTTATTTTGTTCAGA 1680
 QY 1681 AATTTAAAAATAAATTAATGAGCATGGGAAGTTTCACGGGCATCAATTTGAGCAGCACTAGACT 1740
 Db 1681 AATTTAAAAATAAATTAATGAGCATGGGAAGTTTCACGGGCATCAATTTGAGCAGCACTAGACT 1740
 QY 1741 GTTTGAACAATGATGTCCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
 Db 1741 GTTTGAACAATGATGTCCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
 QY 1801 GCATTTCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACCTTTTCTTCAACG 1860
 Db 1801 GCATTTCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACCTTTTCTTCAACG 1860
 QY 1861 GATTGGAATCCTTTTCTAAACTTTTAAAAATAAATAAATGCAATTTTGTAAATTTATC 1920
 Db 1861 GATTGGAATCCTTTTCTAAACTTTTAAAAATAAATAAATGCAATTTTGTAAATTTATC 1920
 QY 1921 AACACCTCAACATGATGTAGCGTACTATAAATAGTGCTCTTGGTGCTCTACTATCAT 1980
 Db 1921 AACACCTCAACATGATGTAGCGTACTATAAATAGTGCTCTTGGTGCTCTACTATCAT 1980
 QY 1981 CACATCAATCTTACACCAACAACTTGCAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
 Db 1981 CACATCAATCTTACACCAACAACTTGCAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
 QY 2041 ATTCTAAATATC 2052
 Db 2041 ATTCTAAAGATC 2052

RESULT 3
 AAA37964

ID AAA37964 standard; DNA; 2056 BP.

AC AAA37964;

XX 18-AUG-2000 (first entry)

XX Plasmid #3 DNA sequence used in mutation of promoter sequence.

XX Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

XX Daucus carota.

XX WO20020613-A1.

XX 13-APR-2000.

XX 28-SEP-1999; 99WO-JP005303.

XX 02-OCT-1998; 98JP-00281124.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX

PI Nishikawa S, Oeda K;
 XX WPI; 2000-303791/26.
 XX
 PT New plant promoters and terminators from Daucus carota L., useful in
 XX plant breeding, for e.g. controlling fertilities of plants.
 PS Example 8; Page 74-76; 81pp; English.
 XX
 CC This sequence represents a plasmid sequence used in a method for
 CC introducing a mutation into a carrot promoter. The invention relates to
 CC plant promoters and terminators from Daucus carota L. which are capable
 CC of expressing a gene of interest in plants. The invention also includes a
 CC chimeric gene characterized in that it comprises the promoter and a
 CC desired gene linked to each other in the form capable of functioning. A
 CC method of producing a transformant comprises introducing the promoter,
 CC the chimeric gene or a vector comprising the promoter and a desired gene
 CC or terminators sequence into a host cell. The plant promoters and
 CC terminators are useful in plant breeding, for e.g. fertilities of plants
 CC may be controlled by expressing, in the host cells, a sense or antisense
 CC gene of a male sterility related gene such as S-locus-specific RNase gene
 XX
 SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;

Query Match 99.2%; Score 2034.8; DB 3; Length 2056;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2050; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 1 CATGTGCGCCTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
 Db 1 CATGTGCGCCTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAGAAGCTGCTTCTGA 60
 QY 61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTCGAA 120
 Db 61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAGTAGAAATATTTTTAAAAAGCTCGAA 120
 QY 121 TACTAATCTCTCTCACAACCTTCCGCTTCTTTTCCAAACACTTTATTACTTTTTTACT 180
 Db 121 TACTAATCTCTCTCACAACCTTCCGCTTCTTTTCCAAACACTTTATTACTTTTTTACT 180
 QY 181 TCTCATTTCTACTCCACTTCTTTGCTATAGCAAGAAATCCTTCTTTAAGCTTAACCCA 240
 Db 181 TCTCATTTCTACTCCACTTCTTTGCTATAGCAAGAAATCCTTCTTTAAGCTTAACCCA 240
 QY 241 AACGGCCTCAATAAAGATCATTCAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
 Db 241 AACGGCCTCAATAAAGATCATTCAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
 QY 301 GAACAGGGTTATTTTTTAAGCTGTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA 360
 Db 301 GAACAGGGTTATTTTTTAAGCTGTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA 360
 QY 361 CCGTCTTCCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCATGC 420
 Db 361 CCGTCTTCCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCATGC 420
 QY 421 AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGACAAGTGGTTC 480
 Db 421 AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGACAAGTGGTTC 480
 QY 481 ATGCTCAGCCATCAAAATTTGACAAAACCCGACACACACTCTATCCACGTACTATCTTT 540
 Db 481 ATGCTCAGCCATCAAAATTTGACAAAACCCGACACACACTCTATCCACGTACTATCTTT 540
 QY 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCATCCAAGGATAAGT 600
 Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCATCCAAGGATAAGT 600
 QY 601 AAAATTTCCCGTTTAAACAGTTTGTATATATATATATGTTTACACTTACAGAGGATATTCGT 660
 Db 601 AAAATTTCCCGTTTAAACAGTTTGTATATATATATATGTTTACACTTACAGAGGATATTCGT 660
 QY 661 AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Matches 2046; Conservative 0; Mismatches 2; Indels 4; Gaps 1;			
Qy	1	CATGTGTCCTCAGACACATAGGCGCTGTTGGTTGAGAGAAGCAGAGCTGTTCTGA	60
Db	1	CATGTGTCCTCAGACACATAGGCGCTGTTGGTTGAGAGAAGCAGAGCTGTTCTGA	60
Qy	61	CTTCTCTCTCTTTTGACCTGTTGTATAAAGAGTGAATAATTTTTAAAAAGCTGGAA	120
Db	61	CTTCTCTCTCTTTTGACCTGTTGTATAAAGAGTGAATAATTTTTAAAAAGCTGGAA	120
Qy	121	TACTAACTCTCTCACAACCTCCGCTCTCTTTCCAAACACCTTTATTAACCTTTTTTACT	180
Db	121	TACTAACTCTCTCACAACCTCCGCTCTCTTTCCAAACACCTTTATTAACCTTTTTTACT	180
Qy	181	TCTCATTTCTACTCCACTCTTTGCTATAAGCAAGAAATCACTCTCTTTAAGCTAACCCA	240
Db	181	TCTCATTTCTACTCCACTCTTTGCTATAAGCAAGAAATCACTCTCTTTAAGCTAACCCA	240
Qy	241	AACGGCTCAATAAAGATCATTAATTAATGTATCTTTCAATTTAGGATAACATACGT	300
Db	241	AACGGCTCAATAAAGATCATTAATTAATGTATCTTTCAATTTAGGATAACATACGT	300
Qy	301	GAACAGGTTATTTTTAAAGTGTCACAAATCTTAATAATTTTACCTGCGCGTGAACA	360
Db	301	GAACAGGTTATTTTTAAAGTGTCACAAATCTTAATAATTTTACCTGCGCGTGAACA	360
Qy	361	CCGTCTTCCAAGATAATATTTTTAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC	420
Db	361	CCGTCTTCCAAGATAATATTTTTAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC	420
Qy	421	AGGACGACTTAGTGAATACACATGTATCTGTAGTCTTTTAAACAAAGACAAGTGGTTC	480
Db	421	AGGACGACTTAGTGAATACACATGTATCTGTAGTCTTTTAAACAAAGACAAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAATTTGAACAAACCCGACACACACTCTATCCAGTACTATCTTT	540
Db	481	ATGCTCAGCCATCAAAATTTGAACAAACCCGACACACACTCTATCCAGTACTATCTTT	540
Qy	541	TGGCCGAATGCTCTCAAAATGTTTTTATATGTAAATAATGCCCCATCCAAGGATAAGT	600
Db	541	TGGCCGAATGCTCTCAAAATGTTTTTATATGTAAATAATGCCCCATCCAAGGATAAGT	600
Qy	601	AAATTCGGTTTAAACAGTTGTTAATATATATGTTTACATTACAAGAGATATTCGT	660
Db	601	AAATTCGGTTTAAACAGTTGTTAATATATATGTTTACATTACAAGAGATATTCGT	660
Qy	661	AATACTTTTAGACACACAGAGACTTAGTCAAAATGGAACGCTGTAACACAGCCTAGACT	720
Db	661	AATACTTTTAGACACACAGAGACTTAGTCAAAATGGAACGCTGTAACACAGCCTAGACT	720
Qy	721	TGGTCACTGATPAAATAGATAATTTAGTATPAAATATAGTAGGATCTACAATGACATPAA	780
Db	721	TGGTCACTGATPAAATAGATAATTTAGTATPAAATATAGTAGGATCTACAATGACATPAA	780
Qy	781	ATTAGAGCTAATTAAGTTACTAATAAATAAGAGAGTTAGTAAACAGAAAGCAGGTA	840
Db	781	ATTAGAGCTAATTAAGTTACTAATAAATAAGAGAGTTAGTAAACAGAAAGCAGGTA	840
Qy	841	AAAAACAAGAGCTGCTGCTGCTGTTAGTTGTTGAGCTCATTTCTTTAAAAAGTAATG	900
Db	841	AAAAACAAGAGCTGCTGCTGCTGTTAGTTGTTGAGCTCATTTCTTTAAAAAGTAATG	900
Qy	901	TAAACTGATCTPAAAGCACAAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA	960
Db	901	TAAACTGATCTPAAAGCACAAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA	960
Qy	961	TTAAACGAATCATTTTATAACATGCTCTCGGCTCATTAATAAGGATCAGCTTAC	1020
Db	961	TTAAACGAATCATTTTATAACATGCTCTCGGCTCATTAATAAGGATCAGCTTAC	1020
Qy	1021	TGATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATAATCTTACAATGAA	1080
Db	1021	TGATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATAATCTTACAATGAA	1080
Qy	1017	TGATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATAATCTTACAATGAA	1076
Db	1017	TGATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATAATCTTACAATGAA	1076

RESULT 5
AAA37963
ID AAA37963 standard; DNA; 2048 BP.

Qy	1081	AAGAAGCAATGCTCTTTTGAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT	1140
Db	1077	AAGAAGCAATGCTCTCTTTTGAAAAACAAATAGGTACTCCCTCCGTCCCTCTGAAATGT	1136
Qy	1141	ATACATATGGATTGGACACGGAGACTTAAGAAAAATGTAATAAGTAATGTAGAGTAAAAAG	1200
Db	1137	ATACATATGGATTGGACACGGAGACTTAAGAAAAATGTAATAAGTAATGTAGAGTAAAAAG	1196
Qy	1201	AAAGAGAAAGAAAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTCATAGATTTAG	1260
Db	1197	AAAGAGAAAGAAAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTCATAGATTTAG	1256
Qy	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTTATATATAATAAAAAATTTACTATTTTG	1320
Db	1257	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTTATATAATAAAAAATTTACTATTTTG	1316
Qy	1321	AGAAAGTTTTGAAATGTAAGAAATTCAGTGGGACATCCATAAAGGAAAGTGTATAGAAT	1380
Db	1317	AGAAAGTTTTGAAATGTAAGAAATTCAGTGGGACATCCATAAAGGAAAGTGTATAGAAT	1376
Qy	1381	TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGATTTCA	1440
Db	1377	TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGATTTCA	1436
Qy	1441	TAAAGATTATAATCTATGTTTATAATGATAATAATAATTTTAAAAATAATATATATTAATT	1500
Db	1437	TAAAGATTATAATCTATGTTTATAATGATAATAATAATTTTAAAAATAATATATATTAATT	1496
Qy	1501	CTGATTAGTCGATTACCGCTTTTATAATTTTACAATCTAGTGAATATGAATAAATCAG	1560
Db	1497	CTGATTAGTCGATTACCGCTTTTATAATTTTACAATCTAGTGAATATGAATAAATCAG	1556
Qy	1561	TTATCTGAAAAAGCAATAATATCTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAGTTCATG	1620
Db	1557	TTATCTGAAAAAGCAATAATATCTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAGTTCATG	1616
Qy	1621	TGATTCAATAGTGTAAATAAAAAAGTAATTTTAAATTTGTTATTTTGTGTTTCA	1680
Db	1617	TGATTCAATAGTGTAAATAAAAAAGTAATTTTAAATTTGTTATTTTGTGTTTCA	1676
Qy	1681	AAATTTAAAAATAATATTCAGCATGGGACATTCACGGGCATCATTTGAGCAGCAGCTAGACT	1740
Db	1677	AAATTTAAAAATAATATTCAGCATGGGACATTCACGGGCATCATTTGAGCAGCAGCTAGACT	1736
Qy	1741	GTTTGAACAATGATGTATGTCGGTGTACATCTATGACCTTTTCAACTAGTGAATAAT	1800
Db	1737	GTTTGAACAATGATGTATGTCGGTGTACATCTATGACCTTTTCAACTAGTGAATAAT	1796
Qy	1801	GCATTTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCAACG	1860
Db	1797	GCATTTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCAACG	1856
Qy	1861	GATTGGAATCCTTTTCTAAACTTTTTTAAAAATAAAAAATGCAATTTGTAATATTTTATC	1920
Db	1857	GATTGGAATCCTTTTCTAAACTTTTTTAAAAATAAAAAATGCAATTTGTAATATTTTATC	1916
Qy	1921	AACACCTCAACATTTAGTGTAGCGTACTATAAAATAGGTGCTCTTGGTCTCTACTATCAT	1980
Db	1917	AACACCTCAACATTTAGTGTAGCGTACTATAAAATAGGTGCTCTTGGTCTCTACTATCAT	1976
Qy	1981	CACATCAATCTTACACCAAAACCTTGAGCTTAAATTTTCTACTTATTTCTCAGCAATAAC	2040
Db	1977	CACATCAATCTTACACCAAAACCTTGAGCTTAAATTTTCTACTTATTTCTCAGCAATAAC	2036
Qy	2041	ATTCTAAATATC 2052	
Db	2037	ATTCTAAAGATC 2048	

Db 1437 TAGATTATAAATCTAGTTATTAATGATAATAATAATTTTAAAAATAATACTATATTAATT 1496
Qy 1501 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACCTGAGTAATATGAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATACCTGAGTAATATGAATAAATCAG 1556
Qy 1561 TTATCTGAAAGCAATAATATCTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTCATG 1620
Db 1557 TTATCTGAAAGCAATAATATCTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTCATG 1616
Qy 1621 TGTATTCAATAGCTTTAATAATAAAGTAATTTTAAATTAATGTTATTTTCTTTTCAGA 1680
Db 1617 TGTATTCAATAGCTTTAATAATAAAGTAATTTTAAATTAATGTTATTTTCTTTTCAGA 1676
Qy 1681 AATTTAAATAAATTTATTGAGCATGGGAAGTTTCACGGGCATCTTTGAGCAGCACTAGACT 1740
Db 1677 AATTTAAATAAATTTATTGAGCATGGGAAGTTTCACGGGCATCTTTGAGCAGCACTAGACT 1736
Qy 1741 GTTTGAAACAATGATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTGAAACAATGATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATCTTAGAATACATCTTTTCAAAATTTTCAACAACACAGCTTTTAACTTTTCTTCAACG 1860
Db 1797 GCATCTTAGAATACATCTTTTCAAAATTTTCAACAACACAGCTTTTAACTTTTCTTCAACG 1856
Qy 1861 GATTGGAATCCTTTTCTTAAACTTTTTTAAAAATAAATAAATGCAATTTATTGTAATATTATC 1920
Db 1857 GATTGGAATCCTTTTCTTAAACTTTTTTAAAAATAAATAAATGCAATTTATTGTAATATTATC 1916
Qy 1921 AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
Db 1917 AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976
Qy 1981 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
Db 1977 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2036
Qy 2041 ATTCATAATATC 2052
Db 2037 ATTCATAAGATC 2048

RESULT 6
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX AAV15144;
AC AAV15144;
XX 02-JUL-1998 (first entry)
XX New promoter used for root-specific expression in plants.
XX Promoter; root; carrot; Kuroda Gosun; root-specific expression;
XX plant cell; soil pathogen; improve; nutritive value; edible root plant;
XX ss.
XX Daucus carota.
XX EP824150-A2.
XX 18-FEB-1998.
XX 12-AUG-1997; 97EP-00113923.
XX 12-AUG-1996; 96JP-00212680.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX Torikai S, Oeda K;
XX WPI; 1998-122310/12.
XX

XX New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX Claim 2; Page 15-16; 31pp; English.
XX The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
SQ
Query Match 11.7%; Score 239.6; DB 2; Length 2042;
Best Local Similarity 72.6%; Pred. No. 8.7e-35;
Matches 366; Conservative 0; Mismatches 129; Indels 9; Gaps 4;
Qy 1554 AAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAACACAGCG---TTCGGTCAAAATG 1609
Db 1543 AAATTTATCTGTAATGATAACATCTTTGTAAACAAACTGCGCCAAATAGGACCATAA 1602
Qy 1610 GGNAGTTTCATGCTATTCAATAGTTTAAATAAAGTAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCAAGTTTCAGTGTATTCTTAAATGTTAATACTAATGATGATTTTCTT--TTCAGGT 1660
Qy 1670 TTTGTTTTCAGAAATTTTAAATAAATTTATGAGCATGGGAAGTTTCACGGGCATCATTTGAGC 1729
Db 1661 ATAGTTAATCTTCAATCAATTAATTTGACATATTGAGCAACTTTATGCC 1720
Qy 1730 AGCACTAGACTGTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAAC 1789
Db 1721 CACGTTGTTATGTTTAAACAAACGTTTGTCCGGTGATATTATGACCTTTCAACTCAAGC 1780
Qy 1790 TAGTGAAT--AATGCATCTTAGAATACATCTTTTCAAAATTTCAACAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAACACAGCACTAACT 1840
Qy 1849 TTTCTTTTCAACGGATTGGAACTCTTTCTTAAACTTTTAAAAATAAATAAATGCAATTTATT 1908
Db 1841 TTTCTTTTAAACAGATTAGAACTCGTTTCGTAAACTTTTAAATTT--AAAAATAACATTA 1898
Qy 1909 GTAATATTATCAACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTATCAACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958
Qy 1969 CTCTACTATCATCAGATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATT 2028
Db 1959 CTCTACTATCATCAGATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATT 2018
Qy 2029 CTCAGCAATAAATCTTCAATAATATC 2052
Db 2019 TTTAGCAAAAACATTTCTAAAGTTC 2042
RESULT 7
AAV15143
ID AAV15143 standard; DNA; 247 BP.
XX AAV15143;
AC AAV15143;
XX 02-JUL-1998 (first entry)
XX New promoter used for root-specific expression in plants.
XX Promoter; root; carrot; Kuroda Gosun; root-specific expression;
XX plant cell; soil pathogen; improve; nutritive value; edible root plant;
XX ss.
XX Daucus carota.
XX

```

FN EP824150-A2.
XX
XX
PD 18-FEB-1998.
XX
XX PF 12-AUG-1997; 97EP-00113923.
XX
XX PR 12-AUG-1996; 96JP-00212680.
XX
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX
XX PI Torikai S, Oeda K;
XX
XX DR WPI; 1998-122310/12.
XX
XX PT New carrot root gene, promoter and terminator - useful in genetic
XX engineering for directing root-specific gene expression.
XX
XX PS Claim 1; Page 14; 31pp; English.
XX
XX CC The present sequence represents a novel promoter, and is isolated from
XX the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
XX direct root-specific expression in plant cells. Since the promoter
XX enables expression of a desired protein in the roots of a plant, it is
XX useful in combat against pathogenic soil fungi and pests which are
XX difficult to kill by chemicals. It can also be used to improve the
XX nutritive value of edible root plants
XX
XX SQ Sequence 247 BP; 88 A; 50 C; 21 G; 88 T; 0 U; 0 Other;

Query Match 9.6%; Score 196.6; DB 2; Length 247;
Best Local Similarity 89.6%; Pred. No. 5.8e-27;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy 1804 TTCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGAT 1863
Db 1 TTCTAGATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACAGAT 60

Qy 1864 TGGAACTCTTTTCTAAACCTTTTAAATAATAAAAAATGCAATATTGTAAATATTATCAAC 1923
Db 61 TAGAATCGTTTCTTAACCTTTTAAATTT--AAAAATACATTAATTAATATTATCAAC 118

Qy 1924 ACCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTTTGGTCTCTACTATCATCAC 1983
Db 119 ACCTCAACATTTCATGTTAGCGTACTATAAAATAGGTGCTTTGGTCTCTACTATCATCAC 178

Qy 1984 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTTATCTCAGCAATAACATT 2043
Db 179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTCTACTAAATTTTTCAGCAAAACATT 238

Qy 2044 CTAATATC 2052
Db 239 CTAAGGTC 247

RESULT 8
AAZ49611
ID AAZ49611 standard; DNA; 246 BP.
XX
XX AC AAZ49611;
XX
XX DT 07-APR-2000 (first entry)
XX
XX DE Carrot CR16.3 fragment for synthesis of plant promoter.
XX
XX KW Synthetic DNA; plant promoter; CR16.3 fragment; carrot; transgenic plant;
XX soybean glycinin; stearyl-ACP-desaturase gene;
XX KW male sterility-related gene; ds.
XX
XX OS Daucus carota.
XX
XX PN EP976832-A2.
XX
XX PD 02-FEB-2000.

XX
XX PF 13-JUL-1999; 99EP-00113732.
XX
XX PR 15-JUL-1998; 98JP-00200372.
XX
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX
XX PI Ishige F, Nishikawa S, Oeda K;
XX
XX XX. WPI; 2000-128374/12.
XX
XX DR Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX
XX PS Claim 1; Page 11-12; 24pp; English.
XX
XX CC The present sequence is a CR16.3 fragment from carrot genomic DNA. It is
XX used for synthesis of a plant promoter which comprises nucleotides 112-
XX 246, 54-246, or 1-246 of this sequence and a synthetic DNA. The promoter
XX is used for controlling the expression of a desired gene e.g. soybean
XX glycinin, stearyl-ACP-desaturase and S-locus type specific RNase gene
XX (male sterility-related gene) in a host cell especially a microorganism
XX or a plant cell. The transformed plant cells can be used to produce
XX transgenic plants. The promoter is compact and therefore suitable for
XX higher expression of a desired gene in a particular tissue compared to
XX other host tissues
XX
XX SQ Sequence 246 BP; 88 A; 50 C; 21 G; 87 T; 0 U; 0 Other;

Query Match 9.5%; Score 195.6; DB 3; Length 246;
Best Local Similarity 89.5%; Pred. No. 8.9e-27;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy 1805 TCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGATT 1864
Db 1 TCTAGATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACAGATT 60

Qy 1865 GGAATCTCTTTTCTAAACCTTTTAAATAATAAAAAATGCAATATTGTAAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACCTTTTAAATTT--AAAAATACATTAATTAATATTATCAACA 118

Qy 1925 CCTCAACATTGATGTTAGCGTACTATAAAATAGGTGCTTTGGTCTCTACTATCATCAC 1984
Db 119 CCTCAACATTTCATGTTAGCGTACTATAAAATAGGTGCTTTGGTCTCTACTATCATCAC 178

Qy 1985 TCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAACATTC 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTTTCTACTAAATTTTTCAGCAAAACATTC 238

Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 9
ADP07499
ID ADP07499 standard; DNA; 196 BP.
XX
XX AC ADP07499;
XX
XX DT 29-JUL-2004 (first entry)
XX
XX DE Carrot DNA.
XX
XX KW Carrot; gene; ds; expression inducing promoter;
XX transcription start point; RNA polymerase II.
XX
XX OS Daucus carota.
XX
XX PN JP2004135597-A.
XX
XX PD 13-MAY-2004.

```

```
PF 18-OCT-2002; 2002JP-00304115.
XX
PR 18-OCT-2002; 2002JP-00304115.
XX
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX WPI; 2004-360986/34.
XX
XX Expression inducing promoter, useful for expressing foreign gene,
XX comprising first DNA for detecting transcription start point and having
XX minimum promoter function, linked to second DNA having expression
XX inducing promoter function.
XX
XX Disclosure; SEQ ID NO 4; 61pp; Japanese.
XX
XX The invention relates to an expression inducing promoter comprising a
XX first DNA linked to a second DNA at its 5' terminus, where the first DNA
XX has a region which determines a transcription start point of RNA
XX polymerase II and has minimum promoter function, and the second DNA has
XX expression inducing promoter function in a plant cell. The expression
XX inducing promoter is useful for expressing a foreign gene. This sequence
XX represents carrot DNA used in the scope of the invention.
XX
XX Sequence 196 BP; 66 A; 40 C; 19 G; 71 T; 0 U; 0 Other;
XX
XX Query Match 9.5%; Score 194.4; DB 12; Length 196;
XX Best Local Similarity 99.5%; Pred. No. 1.4e-26;
XX Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1853 TTTCAACGGATTGGAACTCTTTCTTAACTTTTAAATAAATAAATAATGCAATTATTGTA 1912
Db 1 TTTCAACGGATTGGAACTCTTTCTTAACTTTTAAATAAATAAATAATGCAATTATTGTA 60
XX
QY 1913 TATTATCAACACCTCAACATTTGATTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCT 1972
Db 61 TATTATCAACACCTCAACATTTGATTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCT 120
XX
QY 1973 ACTATCATCATCAATCTTTACACACAAACCTTGAGCTTAATTTTCTACTTATCTCA 2032
Db 121 ACTATCATCATCAATCTTTACACACAAACCTTGAGCTTAATTTTCTACTTATCTCA 180
XX
QY 2033 GCAATACATCTTAAA 2048
Db 181 GCAATACATCTTAAA 196
XX
RESULT 10
ADC56759
ID ADC56759 standard; DNA; 2831 BP.
XX
AC ADC56759;
XX
XX 18-DEC-2003 (first entry)
XX
XX Carrot DNA that encodes an expression inducer type promoter.
XX
XX carrot; expression inducer type promoter; plant; vector; transformant;
XX db.
XX
XX Daucus carota subsp. sativus.
XX
XX JP2003000252-A.
XX
XX 07-JAN-2003.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX WPI; 2003-472920/45.
XX
XX This invention relates to carrot DNA used as an expression inducer type
XX promoter. Specifically, it refers to promoter sequences derived from
XX plant DNA, preferably carrot, and functional mutants thereof that can be
XX
```

```
PT An expression inducer type promoter derived from DNA of Daucus carota L.
var. sativa DC. of 2831 bases and its analogues having the same
functions.
XX
XX Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
XX This invention relates to carrot DNA used as an expression inducer type
XX promoter. Specifically, it refers to promoter sequences derived from
XX plant DNA, preferably carrot, and functional mutants thereof that can be
XX used as expression inducers. Furthermore, the present invention describes
XX DNA, vectors, transformants and the process by which to prepare
XX transformants. This polynucleotide sequence is the DNA encoding the
XX carrot promoter of the invention.
XX
XX Sequence 2831 BP; 950 A; 456 C; 395 G; 1030 T; 0 U; 0 Other;
XX
XX Query Match 7.9%; Score 162.4; DB 10; Length 2831;
XX Best Local Similarity 82.5%; Pred. No. 1.5e-20;
XX Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
XX
QY 22 AGGCGCTGTTGGTTGAGAGAAGCAGAGCTGCTTCTGACTTCTTCTTTGACCTGT 81
Db 1716 AGGCGCTGTTGGTTGAGAGAAGCAGAGCTGCTTCTGACTTCTTCTTTT-ACCGT 1774
XX
QY 82 TTGTATAAAGAGTAGAATATTTTAAAGCTGCGAATACTTAATCTCTCTCACAAC 141
Db 1775 TTGTATAAAGAGTAGAATATTTTAAAGCTGCGAATACTTAATCTCTCTCACAAC 1834
XX
QY 142 TTCGCTTCTTTTCCAAACACTTTATTAACCTTTTAACTTCTTCTTCTTCTTCTTCT 201
Db 1835 TTCGCTTCTTTTCCAAACACTTTATTAACCTTTTAACTTCTTCTTCTTCTTCTTCT 1894
XX
QY 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAACCCAAACGGCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAAATCACTTCTTTTAAAGCTAACCCAAACGGCTCAATAAAGATCA 1954
XX
RESULT 11
ADC56761
ID ADC56761 standard; DNA; 2865 BP.
XX
AC ADC56761;
XX
XX 18-DEC-2003 (first entry)
XX
XX Carrot DNA encoding an expression inducer type promoter (SeqID 3).
XX
XX carrot; expression inducer type promoter; plant; vector; transformant;
XX db.
XX
XX Daucus carota subsp. sativus.
XX
XX JP2003000252-A.
XX
XX 07-JAN-2003.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX
XX 31-MAY-2001; 2001JP-00164069.
XX
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX WPI; 2003-472920/45.
XX
XX An expression inducer type promoter derived from DNA of Daucus carota L.
var. sativa DC. of 2831 bases and its analogues having the same
functions.
XX
XX Disclosure; SEQ ID NO 3; 16pp; Japanese.
XX
XX This invention relates to carrot DNA used as an expression inducer type
XX promoter. Specifically, it refers to promoter sequences derived from
XX plant DNA, preferably carrot, and functional mutants thereof that can be
XX
```

used as expression inducers. Furthermore, the present invention describes DNA, vectors, transformants and the process by which to prepare CC transformants. This polynucleotide sequence is DNA encoding a carrot CC promoter (SeqID 3) of the invention.

SQ	Sequence	2865 BP;	963 A;	461 C;	396 G;	1045 T;	0 U;	0 Other;
	Query Match	7.9%;	Score 162.4;	DB 10;	Length 2865;			
	Best Local Similarity	82.5%;	Pred. No. 1.5e-20;					
	Matches 198; Conservative	0;	Mismatches 41;	Indels 1;	Gaps 1;			
Qy	22 AGGCCCTTTGGTGTGAGACAGACGAAGCTTCCTGCAC TTCCTCCTCTTTTGACCCTGT	81						
Db	1716 AGGCCCTTTGTTTTATGGAATCAGAAGCTGCTCTGCAC TTCCTCCTCTTTT - ACCCGT	1774						
Qy	82 TTGTATAAAGAGTPAGAAATATTTTTAAAAAGCTCGCAATACTAAC TTCCTCTCACAA C	141						
Db	1775 TTGTGTAAAAGAGCAGAAGCACTTTTAAGAAGCTGAGAA TGCTAGTTCTCTCTCACAG C	1834						
Qy	142 TTCGGCTTCCTTTCCAAACA CTATTATAACTTTTTTTACTCTCTCATTTCTACTCCACTTCT	201						
Db	1835 TTCTGCTTCCTTTTCAAACA CTTATCACTTACTCTCTACTTCTCTGCTTCACCTTCT	1894						
Qy	202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAACCCAACGGCCCTCAATAAAGATCA	261						
Db	1895 TTACTATAAGCAAGAGTCAATTCCTTTTAAATTAACCCAACGCCCTCTAGTTAA TTTGA	1954						

RESULT 12	
AAZ49616/c	
ID	AAZ49616 standard; DNA; 140 BP.
XX	
XX	AAZ49616;
XX	
XX	
XX	07-APR-2000 (first entry)
XX	
DE	Oligonucleotide-4 for synthesis of CR16.1 fragment for plant promoter.
XX	
XX	Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
KW	soybean glycinin; oligonucleotide; stearoyl-ACP-desaturase gene;
KW	male sterility-related gene; ss.
XX	
XX	
OS	Daucus carota.
XX	Synthetic.
XX	
PN	EP976832-A2.
XX	
PD	02-FEB-2000.
XX	
XX	
PF	13-JUL-1999; 99EP-00113732.
XX	
XX	
PR	15-JUL-1998; 98JP-00200372.
XX	
XX	(SUMO) SUMITOMO CHEM CO LTD.
PA	
PI	Ishige F, Nishikawa S, Oeda K;
XX	
XX	WPI; 2000-128374/12.
XX	
DR	
PT	Novel promoter used to produce transgenic plants with higher expression
PT	of a desired gene.
XX	
XX	
PS	Disclosure; Page 14; 24pp; English.
XX	
CC	
CC	The present sequence is an oligonucleotide (- chain) used to prepare a
CC	CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
CC	fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
CC	construction of a plant promoter. The promoter is used for controlling
CC	the expression of a desired gene e.g. soybean glycinin, stearoyl-ACP-
CC	desaturase and S-locus type specific RNase gene (male sterility-related
CC	gene) in a host cell especially a microorganism or a plant cell. The
CC	transformed plant cells can be used to produce transgenic plants. The
CC	promoter is compact and therefore suitable for higher expression of a
CC	CC

CC	desired gene in a particular tissue compared to other host tissues	
XX		
SQ	Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;	
	Query Match 5.8%; Score 119; DB 3; Length 140;	
	Best Local Similarity 92.6%; Pred. NO. 1.1e-12;	
	Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	
QY	1918 ATCAACACCTCAACATTGATGTTAGCGCTACTATAAATAGGTGCTCTTTGGTGCTCTACTAT 1977	
Db	140 ATCAACACCTCAACATTGATGTTAGCGCTACTATAAATAGGTGCTCTTTGGTGCTCTACTAT 81	
QY	1978 CATCATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTTATTCAGCAAT 2037	
Db	80 CATCATCAATCTTCCAGCACAAACCTTGAGCTTAATTTTCTACTTATTTTAGCAA 21	
QY	2038 AACATTCTAAATATC 2052	
Db	20 AACATTCTAAAGTC 6	
RESULT 13		
AA249615		
ID	AA249615 standard; DNA; 140 BP.	
XX		
AC	AA249615;	
XX		
DT	07-APR-2000 (first entry)	
XX		
DE	Oligonucleotide-3 for synthesis of CR16.1 fragment for plant promoter.	
XX		
KW	Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;	
KW	soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;	
KW	male sterility-related gene; ss.	
XX		
OS	Daucus carota.	
OS	Synthetic.	
XX		
PN	EP976832-A2.	
XX		
PD	02-FEB-2000.	
XX		
PF	13-JUL-1999; 99EP-00113732.	
XX		
PR	15-JUL-1998; 98JP-00200372.	
XX		
PA	(SUMO) SUMITOMO CHEM CO LTD.	
XX		
PI	Ishige F, Nishikawa S, Oeda K;	
XX		
DR	WPI; 2000-128374/12.	
XX		
PT	Novel promoter used to produce transgenic plants with higher expression	
PT	of a desired gene.	
XX		
PS	Disclosure; Page 13; 24pp; English.	
XX		
CC	The present sequence is an oligonucleotide (+ chain) used to prepare a	
CC	CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3	
CC	fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for	
CC	construction of a plant promoter. The promoter is used for controlling	
CC	the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-	
CC	desaturase and S-locus type specific RNase gene (male sterility-related	
CC	gene) in a host cell especially a microorganism or a plant cell. The	
CC	transformed plant cells can be used to produce transgenic plants. The	
CC	promoter is compact and therefore suitable for higher expression of a	
CC	desired gene in a particular tissue compared to other host tissues	
XX		
SQ	Sequence 140 BP; 44 A; 34 C; 17 G; 45 T; 0 U; 0 Other;	
	Query Match 5.8%; Score 119; DB 3; Length 140;	
	Best Local Similarity 92.6%; Pred. NO. 1.1e-12;	
	Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;	

Qy 1918 ATCACACCTCAACATTTGATGTTAGCGTACTATTAATAGTGCTCTTGGTCTCTACTAT 1977
 Db 5 ATCAACACCTCAACATTTGATGTTAGCGTACTATTAATAGTGCTCTTGGTCTCTACTAT 64

Qy 1978 CATCACATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACATTTCTCAGCAAT 2037
 Db 65 CATCACATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACATTTCTCAGCAAT 124

Qy 2038 AACATCTCTAAATATC 2052
 Db 125 AACATCTCTAAAGTCT 139

RESULT 14
 AB210246
 ID AB210246 standard; DNA; 8056 BP.
 AC AB210246;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN W020027272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-EP003401.
 XX
 PR 26-MAR-2001; 2001US-0278333P.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Mater S, Model P, Mueller V, Otto T, Pelet C;
 PI Schwobe I, Ziebarth H;
 XX
 DR WPI; 2003-019942/01.
 XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX
 PS Claim 28; SEQ ID NO 386; 117pp; English.
 XX

The present invention describes a method for detecting and
 differentiating between hematopoietic cell proliferative disorders
 associated with at least 1 gene and/or their regulatory regions in a
 subject. The method comprises contacting a target nucleic acid in a
 biological sample obtained from the subject with at least 1 reagent,
 which distinguishes between methylated and non-methylated CpG
 dinucleotides within the target nucleic acid. AB209861 to AB21118
 represent specifically claimed nucleotide sequences from the present
 invention. Oligonucleotides from the present invention can be used: for
 differentiating between healthy hematopoietic cells and proliferative
 disorder hematopoietic cells; for differentiating between acute
 lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 determining the cytosine methylation state and/or single nucleotide
 polymorphisms (SNPs) of hematopoietic cell proliferation disorder
 related sequences and their complements; and as primers for the
 amplification of hematopoietic cell proliferation disorder related DNA
 sequences. The nucleotide sequences from the present invention can also
 be used for detecting a predisposition to, differentiation between
 subclasses, diagnosis, prognosis, treatment and/or monitoring of
 haematopoietic cell proliferative disorders. The present method enables a

CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX
 SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
 Query Match 4.8%; Score 99.4; DB 8; Length 8056;
 Best Local Similarity 44.5%; Pred. No. 7.2e-09;
 Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;

Qy 548 ATGCTTCTCAAAATGTTTTTATATGTAATAATGCCATCCCAAGGATAGTAAATTC 607
 Db 1975 ATTAATATTAATTTTATTTAAATTTAAATAATTTAAATATATTTAAATATATAT 2034

Qy 608 CGGTTTAAACAGTTTGTATATATATATGTTTACACTTACAGAGGATATTCGTAATACTT 667
 Db 2035 ATTTTAAATATTTTATTTTAAATAATTTTATTTAAATAATTTTAAATATATATAT 2094

Qy 668 TTAGACGACAGAGACTTAGGTCAAAATGCGCTGGTAAACAGCCTAGACTTGGTCAC 727
 Db 2095 AATAATTTAATAATTTAATAATAATTTTAAATATTTTATTTTAAATAATAAAAA 2154

Qy 728 TGATAAATAGATAAATGTTTAGTATAATATAGTAGGATCTACAATGACATTTAAATTAGAG 787
 Db 2155 TAAAAAATAAAAAAATTTATAATTTTAAATAATTTAAATAATTTAAATAATTTAA 2214

Qy 788 CTATTAATTAAGTTTACTTAATAATAAGAGAGGTTAGTAAACAGACAGGTAACAA 847
 Db 2215 AAAAAAATAAATTTAAATTTTAAATAATTTTATTTTAAATAATTTAAATAATTTAA 2274

Qy 848 GAGCTTGCCTGCTGTGTTTAGTGTGTGAGCTCAATTC-----TTTAAAGTAAATGT 901
 Db 2275 AATTTTAAATTTTAAATTTTATTTTAAATAATTTTAAATAATTTTGTGTTTAAATA 2334

Qy 902 AAACCTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACATTTTACAGAATTTTATAT 961
 Db 2335 AATTTTATTTTAAATAATAATTTTATTTTAAATAATTTTAAATAATTTTATTTAT 2394

Qy 962 TAAACGAATAATCATTTTATTAACATGCTCTCGGCTGCATTTAATAGGATCACITACT 1021
 Db 2395 TATTTTGTGAAATTTAAATAATAATTTTATTTTAAATAATTTAAATTTTATTTT 2451

Qy 1022 GATCATCCATTAAACACCTGTGTAAACAAATTCATAGAGATAAAATATCTTACAATGAAA 1081
 Db 2452 TTATTTTATTTATATATTTTATTTTATTTTGTGTTTAAATAATTTAAATAATTTA 2511

Qy 1082 AGAAGGACATGTCTCTTTGAAAAAACAATAGGTACTCCCTCGCTCCCTGTAAGTGA 1141
 Db 2512 AATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2571

Qy 1142 TACATATGGATTGGACACGAGACTAAGAAATAATGTATAAGTAATGTAGAGTAAAAAGA 1201
 Db 2572 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2631

Qy 1202 AAGAGAAAGAAAAAGTGGTAAAGTAGCGGACCCACCAATATATATATATAGATTTAGA 1261
 Db 2632 AAAAAATAAAAAAATAATTTAAATAATTTAAATAATTTAAATAATTTTAAATTTT 2691

Qy 1262 AAAGTAGTTGAAAGTAGTGGTGGGATTTTATATTTATTTAAATAATTTTACTATTTGA 1321
 Db 2692 TAAATTTATTTATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 2747

Qy 1322 GAAAGTTTGAATGTATAGAAATTTAGTGGGACATCCATAAAGGAAGTGTAGAAAT 1381
 Db 2748 AAAAAATAAAAAAATAATTTAAATTTTATTTTAAATAAATAAATAAATAAATAAATA 2807

Qy 1382 AAATGGACAGAGGAGTAACTTTATGATATATAAATTTTGTATTTTGTATTTTCAAT 1441
 Db 2808 AATTTGTTTATTAATAATAATGATATTTT--TTTATAAATAATTTAAATAATTT 2865

Qy 1442 AAGATTATAAATCTATGTTTATAATGATAATAATAATTTTAAATAATAATACTATTAATTC 1501
 Db 2866 TTTTATGAAAAAATTTTGTGTGAAAAAATAAATAAATAAATAAATAAATAAATAAATA 2925

Qy	1502	TGATTAGTCGATTACCGCCTTTTATAATTTTACAACTAGTAAATGAATAAATCACT	1561
Db	2926	AAATTTTATGAAATGTAAAAATTTAAATGAATTTATTTGTTAAAAAAATAAAAAAATTA	2985
Qy	1562	TATCTGAAAAAGCAAAATAATCTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAGTTCAATGT	1621
Db	2986	TTGAAAAAATTCGATATATTTTAAAAATGATGATGTTTAAATTTTATAAATATAAATAA	3045
Qy	1622	GTATTCAAATAGTTTTTAATATATAAAAGTAAATTTTAAATTAATGTATTTTGTTCGAA	1681
Db	3046	TAAAAATAAATGTTTATAAATGTGAATGAAATTTGAAATTTATTTGTTATTTTATGTCGTAATTT	3105
Qy	1682	ATTTTAAAAATAAAATTTATTGAGCATCGGGAAGTTTCAGGGCATCATTTGACGACGACCTAGACTG	1741
Db	3106	AAATAAAAAAATTAATTTTTTTTAAATGAATTTAATGATGATTGAAAAAATTTGTTTTATATAT	3165
Qy	1742	---TTTGAACAATGTATGTCCGGGTGCATCTATGACCTTTTCAACTCAAACTAGTGAATA	1798
Db	3166	AAATTTTAAATAATAATATTTTAAATAAAAAATTTTGTGTTTGAATTCGTAAATAAAATTTT	3225
Qy	1799	ATGCATCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCCTTTTAACTTTTCTTTCAA	1858
Db	3226	AAATTTATATATATTTTTTTTTTAAATAAATTTGATTAATTTTTTAAATTAATGATTTGTAAAAA	3285
Qy	1859	CGGATTTGGAATCCCTTTTCTAAACCTTTTAAAAATAAAAAAAATGCATTTATTGTAATATTTTA	1918
Db	3286	AAATTTTGTATTGTGTGTTGTTATTTTTTGAAAAAAATAAATAAATTTGTAATGTAAA	3345
Qy	1919	TCAACACCTCAACATTCGATGTTA	1941
Db	3346	AAAATGAATAAAATTTTGGAAATGA	3368

RESULT 15	
ABZ10100	
ID	ABZ10100 standard; DNA; 8056 BP.
XX	
AC	ABZ10100;
XX	
DT	16-JAN-2003 (first entry)
XX	
DE	Haematopoietic cell proliferation disorder related DNA sequence #240.
XX	
KW	Human; haematopoietic cell proliferation disorder; cytostatic;
KW	gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW	cytosine methylation state; gene; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO20027272-A2.
XX	
PD	03-OCT-2002.
XX	
PF	26-MAR-2002; 2002WO-EP003401.
XX	
PR	26-MAR-2001; 2001US-0278333P.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI	Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI	Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI	Schwöpe I, Ziebarth H;
XX	
DR	WPI; 2003-018942/01.
XX	
PT	Detecting and differentiating between hematopoietic cell proliferative
PT	disorders, comprises contacting a target nucleic acid with a reagent that
PT	distinguishes between methylated and non-methylated CpG dinucleotides.
XX	
PS	Claim 28; SEQ ID NO 240; 117pp; English.
XX	
CC	The present invention describes a method for detecting and

CC	differentiating between haematopoietic cell proliferative disorders
CC	associated with at least 1 gene and/or their regulatory regions in a
CC	subject. The method comprises contacting a target nucleic acid in a
CC	biological sample obtained from the subject with at least 1 reagent,
CC	which distinguishes between methylated and non-methylated CpG
CC	dinucleotides within the target nucleic acid. AB209861 to AB211118
CC	represent specifically claimed nucleotide sequences from the present
CC	invention. Oligonucleotides from the present invention can be used: for
CC	differentiating between healthy haematopoietic cells and proliferative
CC	disorder haematopoietic cells; for differentiating between acute
CC	lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC	determining the cytosine methylation state and/or single nucleotide
CC	polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC	related sequences and their complements; and as primers for the
CC	amplification of haematopoietic cell proliferation disorder related DNA
CC	sequences. The nucleotide sequences from the present invention can also
CC	be used for detecting a predisposition to, differentiation between
CC	subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC	haematopoietic cell proliferative disorders. The present method enables a
CC	highly specific classification of haematopoietic cell proliferative
CC	disorders allowing for improved and informed treatment of patients
XX	
SQ	Sequence 8056 BP; 3711 A; 371 C; 371 G; 3503 T; 0 U; 0 Other;
Query Match 4.5%; Score 91.8; DB 8; Length 8056;	
Best Local Similarity 44.2%; Pred. No. 1.8e-07;	
Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;	
Qy	547 AATGCTTCTCAAAATGTTTTTATATGTAAATAATGCCCATCCAGGATGAAGTAAATTT 606
Db	
Qy	1358 ATTTTTTATTAAAAATTTATTTTTTTTTTATAAAAAATAAATATCGATTTTTTTTTTATTTT 1417
Db	
Qy	607 CCGGTTTAAACGATGCTTAAATATATATGTTTACACTTACAAGAGGATATTCGTAATACT 666
Db	
Qy	1418 TTATTTTTTTTTTTAAAAAAAATAAAAAATTTATTTTTTAATAAATTTATTTAAA 1477
Db	
Qy	667 TTTAGACGACAAGAGACTTAGGTCAAAAATGGAGCGTGGTAAACAGCGCTAGACTGGTCA 726
Db	
Qy	1478 ATAATTAATTTAAATTTAATATATATAAATAATAAAATATAAATACGTTAAAAAAATTT 1537
Db	
Qy	727 CTGATAAATAGATAATTGTTAGTATAATATAGTAGGATCTACATGACATTAATAATTAGA 786
Db	
Qy	1538 TTAATAAATAATTTATTTTTTTAAAAAAATATATTTAAATTTTAATTTATATATTTA 1597
Db	
Qy	787 GCTATTAAATTAAGTTTACTTAATAATAAGACAGGTTAGTAAACAGAACGAGGTAAAAACA 846
Db	
Qy	1598 TAATTTTAAATTTTAAAAAATCGAAATAAACGATCGTAAAAATTAATAAATAATTTT 1657
Db	
Qy	847 AGAGCTTGGCTGCTGTGTTTAGTTGTGTGAGCTCATTTCTTTTAAAGTAATGTAAACT 906
Db	
Qy	1658 AATTAATAC-GATAAAATTTTATTTTTTATTAATAATATAAAATTAATTTAATAATAA 1716
Db	
Qy	907 GATCTTAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAGAAATTTTATTTAAAC 966
Db	
Qy	1717 TATATATTTATAATTTTAATTTATTTAAAAAATTTAAAAATAAAAAAATTTTTTAAATTTAAAA 1776
Db	
Qy	967 GAAAAATCATTTTATAACATGCTCTCGGCTGTCATTTATATATAGGATCACTTACTGTATCA 1026
Db	
Qy	1777 AATTAATAAATTAATACGTTTATTTATTTTAAATTAATAAATAAATAAATTTTCGGTTTA 1836
Db	
Qy	1027 TCCATTTAAAAACCTTGTGTAAAAACAAATTCATAGAGTAAAAATATCTTACAAATGAAAGAAG 1086
Db	
Qy	1837 AAAAAAATAAAAAAATAAAAAAATAAATTTAAATTAATTTATTTTATTTTATTTAT 1896
Db	
Qy	1087 GACAATGCTCTTTTGAAAAAATAAATAGGTACTCCCTCCGCTCGAATGTATACAT 1146
Db	
Qy	1897 AAAAAATAAAAAATTTTATAAAAAAATAAATTTAAAAATTTAATAATAAATAAAAAAATAT 1956
Db	
Qy	1147 ATGGATTTGACCGGAGACTTAAGAAAAATGATAAAGTAAATGTAGAGTAAAAAGAAGAG 1206
Db	
Qy	1957 AATTTTTTAAAAATAAATAATAATTAATTAATTTTATTTAAAAATTTAATAAATTTTAAAT 2016
Db	
Qy	1207 A-----AAGAAAAGTGGGTAAAGTAGAGCGGACCCCAACCAATATATTAATTTGATAGAT 1256

Db	2017	ATATTTTAAAAATAAAATATATTTTTTAAATATTTTTTATTTTTTAAAAAATTTTTTATTTAAAAAT	2076
Qy	1257	TTAGAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATTAATAAAAAATTTTACTAT	1316
Db	2077	ATTTTAAATATTATAAATAAATAATTAATAATATTATTAATATAATTTTAAATTTTTTAT	2136
Qy	1317	TTTGAGAAAGTTTTGAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATA	1376
Db	2137	TTTTATTAAAAATAAAAAATAAAAAAATAAAAAATTTATAATTTTTATTAAAAAATTTAAAAAA	2196
Qy	1377	GAATTAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTTCGTATTTTGAT	1436
Db	2197	ACGAAAAAATAAATAAAAAAATAAAAAATTTTAAATTTTAAATAAATAATTTTAAATA	2256
Qy	1437	TTCCATAAGATTATAAATCTATGTTTATTAATGATATAATATAATTTTTTAAAAAATAACTATATT	1496
Db	2257	TTTTAAAAATAAATAAAAAAATTTTAAATTTTTTAAATTTATTTTTTAAAAATAATTTATT	2316
Qy	1497	AAATTCIGATTAGTCGATTACGGCTTTTATAATTTTACAACTACTGAGTATATGAATAAA	1556
Db	2317	AAATTTTCGTAATTAAATAAATAATTTTATTAAATTTTAAATATAATTTTTTAAATTTTATAAA	2376
Qy	1557	TCAGTTATCTGAAAAAGCAAT-----AATATCTTTGTAAAAACAGCGCTTCGGTCAAAATGG	1610
Db	2377	TAAATTTATAAATTTTATTTATTTTTCGAAATTTAAAAAATAAATAATTTTATTTTATA	2436
Qy	1611	GAGTTTCATGTGTATTCGAATAGTTTTTAATATAAAGTAAATTTTAAATTTT-----AATGTGT	1666
Db	2437	AAAAATTAATTTTTTTTTTATTTTATTTATATATTTTTTATTTATTTTCGTTTTTAAAAATAAT	2496
Qy	1667	ATTTTTCGTTTCAGAAATTTAAAAATAAATATTGAGCATGGGAAGTTTCACGGGCATCATTG	1726
Db	2497	AAAAATATAAGAAAAATAAAAAAATAAATAAATTTAATAATAAATAATTTAAAAAATTTTA	2556
Qy	1727	AGCAGCACTAGACTGTTTGAACATGTATGTTCGGGTGACATCTATGACCTTTCAACTCA	1786
Db	2557	AAAAATTTAAATAAATTTTTTAAAAAATAAAAAATAAACGAAAAAATAATATAAATAATA	2616
Qy	1787	AACTAGTGAATAATGCATCTAGAAATACATCTTTTCAAAATTTCAACAAAAACACAGCTTTAA	1846
Db	2617	AATAAATTTAAAAAATAAAAAATAAAAAAATAAAAAAATTTTAAAAATTTTAAAAAATAATA	2676
Qy	1847	CTTTTCTTCTCAACGGATTGGAATCCTTTTCTAAAACTTTTTTAAAAATAAAAAAATGCAATTA	1906
Db	2677	AT-----TTTTTAAATTTTTTAAATTTTATTTTTTTTTTTTTTAAATAAATAATTCGTTA	2732
Qy	1907	TTGTAATATTATCAACACCTTCAACATGTGATTTAGCGT	1945
Db	2733	AATAAATTTATTAATAAAAAATAAAAAAATAATTTTAAATCGT	2771

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2005, 15:57:02 ; Search time 6848 Seconds
(without alignments)
11405.941 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtgcctacagcaca.....gcaataacattctaataatc 2052

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: gb_est1.*

2: gb_est2.*

3: gb_hc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gsl1.*

9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	97.8	4.8	1101	9	CNS00EVL
C 2	97.2	4.7	1101	9	CNS0039G
C 3	95.6	4.7	1542	9	AG386981
C 4	95	4.6	1101	9	CNS0039G
C 5	94.6	4.6	1758	9	CL509408
C 6	93.6	4.6	1896	9	CG753083
C 7	93	4.5	1202	8	CC262481
C 8	92.4	4.5	1101	9	CNS0021J
C 9	90.8	4.4	1101	9	CNS00EVL
C 10	90.4	4.4	1608	9	CL118721
C 11	88.8	4.3	1067	6	CD386564
C 12	88.8	4.3	1268	9	AG347098
C 13	88.8	4.3	1392	9	CG757503
C 14	88	4.3	1277	8	CC253231
C 15	88	4.3	1696	9	AG346840
C 16	87.6	4.3	1275	9	CL033318
C 17	87.2	4.2	887	9	AG526041
C 18	87	4.2	1275	9	CL033318
C 19	87	4.2	2087	9	AG333887
C 20	86.8	4.2	1348	9	CG749499
C 21	85.6	4.2	1350	9	CL019486
C 22	85.4	4.2	1780	9	AG320553
C 23	85.2	4.2	1320	9	CL103881
C 24	85	4.1	1539	9	AG340947

C 25	84.8	4.1	1254	9	AG349719
C 26	84.6	4.1	987	9	CNS014PQ
C 27	84.6	4.1	1364	9	CG757970
C 28	84	4.1	1101	9	CNS017KE
C 29	83.8	4.1	822	7	CK416977
C 30	83.6	4.1	1489	9	AG350139
C 31	83.4	4.1	1101	9	CNS00FMC
C 32	83	4.0	1211	9	AG349657
C 33	83	4.0	1592	9	CG750135
C 34	82.8	4.0	1210	9	CG749728
C 35	82.6	4.0	1506	9	AG278469
C 36	81.6	4.0	1092	9	CNS020K7
C 37	81.6	4.0	1101	9	CNS00EO7
C 38	81.6	4.0	1355	9	AG346348
C 39	81.6	4.0	1745	9	AG338221
C 40	81.4	4.0	812	8	BH178455
C 41	81.4	4.0	812	9	CNS07KRM
C 42	81.4	4.0	1050	9	CNS013NS
C 43	81.4	4.0	1391	9	CG754863
C 44	81.2	4.0	1094	9	CNS012FZ
C 45	81.2	4.0	1715	9	AG288305

ALIGNMENTS

RESULT 1
CNS00EVL/c
LOCUS
DEFINITION
CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL069706
AL069706.1 GI:4949849
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Phyllozoa; Ephydrozoa; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match 4.8%; Score 97.8; DB 9; Length 1101;
Best Local Similarity 34.8%; Pred. No. 8.7e-08;

AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 TITLE BAC end Sequences of Library MSMg01
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1542)
 AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 JOURNAL Direct Submission
 TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@psc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
 library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
 Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1
 Koyadai, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp
 PRIMERS
 Sequencing : TJ
 LIBRARY Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI
 FEATURES
 source Location/Qualifiers
 1. 1542
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSMg01-201G10.TJ"
 /sex="male"
 /tissue type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"
 ORIGIN
 Query Match 4.7%; Score 95.6; DB 9; Length 1542;
 Best Local Similarity 44.3%; Pred. No. 2.2e-07;
 Matches 420; Conservative 0; Mismatches 523; Indels 6; Gaps 1;
 Qy 750 ATATATAGTAGGATCTACATGACATTAAATTAAGAGCTATTAAATTAAGTTACTATAAA 809
 Db 1103 ATAAAGAAATTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1044
 Qy 810 ATAAAGAGAGGTTAGTAAACAGACAGAGGTAACCAAGAGCTGCTGCTGCTGTTAG 869
 Db 1043 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 984
 Qy 870 TTGTTGTGAGCTCATTTCTTTAAAGTAATGTAACCTGATCTAAAGCACATAGAAATTTA 929
 Db 983 TATAATAATATATATAATAATTAATAATAATAATAATAATAATAATAATAATAATAATA 924
 Qy 930 GTACAGGTTAAACCTTTTCAAGAAATTTATATTAAACGAAATCATTTTATAACATGTCT 989
 Db 923 ATAAATAATAAAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 864
 Qy 990 CTCGGCTGCTATTATATAGGATCACTTACTGATCATCCATTA-----AACCTTGTT 1043
 Db 863 TTAATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 804
 Qy 1044 AAAACAAATTCATATGAGATAAATAATCTTCAACATGAAAAGAGCAATGCTCTCTTGAA 1103
 Db 803 AAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 744
 Qy 1104 AAAACAAATAGGTTACTCCCTCGCTCTCTGAAATGTATATACATATGGAATGGACCGAG 1163
 Db 743 TAAATAAATAATATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 684
 Qy 1164 ACTAAGAAAATGTATAAAGTAATGTAGGTAAAGAAAGAGAGAGAAAGTGGGTAAA 1223
 Db 683 AATAATAATAATATTTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 624
 Qy 1224 GTAGCGGACCCCAACCAATATATAATTTGATAGATTTAGAAAAAGTAGTTGAAAGTAGTGGGT 1283

Db 623 ATAAAAAATAAAAAATTTTAAATATTTTAAATNTAAATTTTAAAAAATAAAAAATATAT 564
 Qy 1284 GGGTGGGATTTTATATATATAAAAAATTTACTATTTTCAGAAAGTTTTCGAATGTATAGAA 1343
 Db 563 TATATATAAAAAATATAATAATAAAAAATATAAAAAATATAAAAAATATAAAAAATATA 504
 Qy 1344 TTGAGTGGGACATCCATATAAAGGAAGTGTATAGAAATTAATGGACAGAGGGAGTATAA 1403
 Db 503 TAAAAATATATAAAAAATATTTTATTTTATTTATATATATAAAAAAATAATTTTATAAAA 444
 Qy 1404 CTTTATGATATATAAAATTTTGTATTTTGTATTTTCAATATAGATATATAAATCTATGTTATA 1463
 Db 443 ATTATAAATAAAAAAATTTTATAAATATATAAAAAAATAAATTTTATATTAATAATATA 384
 Qy 1464 ATGATATATATAATTTTAAAAATATACTATATTAATTTCTGATTAGTCGATTACCGCCTTT 1523
 Db 383 ATAAAAAATAAAAAATATAAAAAATTTTATTAATAATTTTAAATAAATAAATAAATAAAT 324
 Qy 1524 TATAATTTTACAATCTGAGTAAATATGAATAAAATCAGTTATCTCGAAAGCAAAATAATC 1583
 Db 323 ATATATATATAAAAAATAAATTTTATTAATAAATAAATAAATAAATAAATAAATAAATA 264
 Qy 1584 TTTGTAAACAGCGTTCGGTCAAAATGGAAGTTCATGTTCAATGTTTCAATGTTTATAATA 1643
 Db 263 AATATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 204
 Qy 1644 AAGTAAATTTTAAATTAATTTGTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1692
 Db 203 AAAATATATATAATTTAAAAAATTTTTTAAATAATTAATAAATAAATAAATAAATAA 155
 RESULT 4
 CNS0039G
 LOCUS
 DEFINITION Drosophila melanogaster genome survey sequence TEF3 end of BAC #
 BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL063921.1 GI:4941778
 VERSION
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mamoser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 FEATURES
 source Location/Qualifiers
 1. 1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR08K10"

```

/clone_lib="RPCI-98"
/notes="end : TET3"

ORIGIN
Query Match      4.6%; Score 95; DB 9; Length 1101;
Best Local Similarity 20.2%; Pred. No. 2.8e-07;
Matches 173; Conservative 325; Mismatches 348; Indels 9; Gaps 2;

Qy 1200 GAAAGAGAAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATATGATGATTTA 1259
Db 189 GTWAKGGGGAATATARGGGGGGAARGGGGGGAARARGRGRGRRAAARAAA 248
Qy 1260 GAAAGTAGTGAAGTAGTGGGTGGGTGGATTTTATATATATATATATATATTTT 1319
Db 249 AARGRRARRGGRRGRGRGRGRGGGGGNGGNGGARAARRRRGRARAGRRRRGGNGN 308
Qy 1320 GAGAAAGCTTTTGAATCTATAGAAATTCAGTGGGATCCATATAAAGGAAAGTGATAGAA 1379
Db 309 GARGRRRRRRRRRAAAGGRRRRRRRAGGGRRRRRRRRRRRRRRRRRRRGARGAGG 368
Qy 1380 TTAATCGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGAATTC 1439
Db 369 RRGAGGGGRRGGGGGGGGGGMATATAWAWWWTTTTTTTTTAWAAWAAATAA 428
Qy 1440 ATAAGATATAAATCTATGTTATATGATATATAATTTTAAATAATATCTATTAAT 1499
Db 429 TTWAAWAAAAAATTTWAAAAWAAAAWATWTTTATWAAAAAATAAATTTT 488
Qy 1500 TCTGATTAGTCGATCCGGCTTTTATAATTTTACATATCTGATATATGAATAATCA 1559
Db 489 TTTTWTWTATTTWATAWTTTWTWTTAAWAAAAAATAAATAAATAATATWTTWT 548
Qy 1560 GTTATCTGAAAGCAAAATATATCTTTGTAAGACAGCGTTCGGTCAATCGGAAGTTTCA 1619
Db 549 WWTYTWTWAAWATAAAMCAAWYHTTYYHYTYTWTYTWYTHWYHWTYTHAWHTWY 608
Qy 1620 GTGATTTCAATAGTTTAAATATAAAGTAATTTTAAATTAATTTGTTATTTTGTTCAG 1679
Db 609 HWYHTYAMWMTWHTWMAWMTTWTYTAAYYYTYTCWY-----YHYMHHHAHAHA 662
Qy 1680 AATTTAAATAAATTTATGAGCATGGGAAGTTTCACGGGCATCATTCGACGACCTAGAC 1739
Db 663 AWTHTTWTWTHAYHWATYHYHYMYCAMCMCHTCHCHYHYHYHTHHTHTHTWYAHY 722
Qy 1740 TGTTTGAACATGATGTCCGGGTACATCTATGACCTTTCAACTCAAACTAGTGAATAA 1799
Db 723 YMWYMYWYMYMYCTACTYVHHHHHYHWAHYHTWYAWAHMMWHHAHYAAAAWA 782
Qy 1800 T---GCATTCAGAAATACATCTTTTCAAAATTCACAAACACAGCTTTAATCTTTCTTC 1856
Db 783 TTHHHHTTHYMHHTYMYHYMYTCCYMCYVCHWHYHTATCTWTHWMTWTHWYHH 842
Qy 1857 AACGGATTGGATCTTTCTTAACTTTTAAATAAATAAATAAATGCAATTTGTAATTT 1916
Db 843 TWHHHTTTHWAWHTHTWCMWHTTWTWATHCWACMTWHHWHHWHHWHHWHHWHH 902
Qy 1917 TATCAACACCTCAACATTTAGGTGATCTATAAATAGTGTCTTTGGTGTCTCTACTA 1976
Db 903 THMCCHHHHCHCTCHHHHTYHMTCHWMMHWHHWHHWHHWHHWHHWHHWHHWHH 962
Qy 1977 TCATCAGATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAA 2036
Db 963 CHMYHMHMYMYCCHYCTCTHTTHATYHMYCTCYHCTWHTYWTYWAWTAHAMTTAT 1022
Qy 2037 TAACATTTCAATAT 2051
Db 1023 WWWHHWAWHATWW 1037

RESULT 5
CL509408
LOCUS
DEFINITION SAIL_811_H11.v3 SAIL Collection Arabidopsis thaliana genomic clone

SAIL_811_H11.v3, genomic survey sequence.
CL509408
VERSION
CL509408.1 GI:46006728
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1758)
Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,
Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D.,
Bullis, D., Snell, J., Miguel, I., Hutchison, D., Kimmerly, B.,
Mitze, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
22356987
PUBMED
12468722
COMMENT
Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwalis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS936276; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1..1758
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_811_H11.v3"
/clone_lib="SAIL Collection"
/notes="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"

FEATURES
source
Query Match      4.6%; Score 94.6; DB 9; Length 1758;
Best Local Similarity 34.2%; Pred. No. 3.3e-07;
Matches 502; Conservative 0; Mismatches 957; Indels 7; Gaps 5;

Qy 495 AAATTGCAAAACCCGACACACACTCTATCCACGTACTACTTTTGGCGGAATGCTTC 554
Db 68 ANANNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTT 127
Qy 555 TCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAGTAATAATCCCGTTTA 614
Db 128 NNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 187
Qy 615 ACCAGTTTGTATATATATGTTTACACTTACAGAGGATATTCGTAATACTTTAGACG 674
Db 188 AANNNNNNNNNNNAANAAAAAANNAANAAAAAANAAAAAANAAAAAANAAAAAN 247
Qy 675 ACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCCTAGACTTTGGTCACTGATA 734
Db 248 ANAAAAANNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAN 307
Qy 735 TAGATAATGTTTGTATATATATAGTAGGATCTCAATGACATTAATAATTAGAGTATTA 794
Db 308 NNNAAAAAANNAANNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 367
Qy 795 TTAAGTTTACTAATAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAACAAAGAGCTTG 854
Db 368 TTTTATTTTANTAAANAAAAAANAAAAAANAAAAAANNNNNNNNNNNNNNNNNNN 427
Qy 855 CTGCTGTGTGTTTGTGTTGTTGAGCTCATTTCTTTAAAAAGTAAATGTAATCTGATA 914
Db 428 NNNNAANNAANAAAAAANAAAAAANNAANNAANNAANNAANNAANNAANNNNNNTANA 487
```



```

Db      1413 TAAATTTTATTTTAAATATTTAAATATATTTAAATTTNANNAAATATAAATATATAT 1354
Qy      1010 GGATCACTTACTGATCATCCATTAACACCTTGTGTAAACAAATTCATAGAGATAAAATAT 1069
Db      1353 AAAAAATATNAATTAATAAATAATTAATTTTAAATTTTAAATTTTAAATTTAAATTTAT 1294
Qy      1070 CTTACATGAAGAAGGACATGCTCTTTTGAACAAACAAATAGGTACTCCCTCCGTC 1129
Db      1293 AAAAAATAAAAAATAAAAAATTTATATAAATAAATAATATATATNAAAAAAATAAATAA 1234
Qy      1130 CTCTGAAATGTATACATATCGATTGGACACGGAGCTTAAGAAAAATCTATAAAGTAAATGT 1189
Db      1233 AAAAAATATATATATTAATATATATATAAATAAATAAATTAATTAATAAATAAATTTATTT 1174
Qy      1190 AGAGTAAAAAGAAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCCAACCAATATATAAT 1249
Db      1173 ATAAATATAATATAAATAAATAAATAAATAAATAA-----AAAAAATAAATAT 1123
Qy      1250 GATAGATTTAGAAAGTAGTTGAAGTAGTGGGTGGGTGGGTGGGTATTTTATATTTATAAAT 1309
Db      1122 TATAAATTTAAAAATAAAAAATNTATTTTAAATTTAATAAATAAATAAATTTATTTAT 1063
Qy      1310 TTACTATTTTGAGAAAGTTTGGAAATGTATAGAAATTCAGTGGGACATCCATAAAGGAAA 1369
Db      1062 TTATATAATATTAATTAATTTNTTTTATATATATATATATAAATAAATTTTATAAAT 1003
Qy      1370 GTGTATAGAATTAATGGGACAGAGGAGTAATACCTTTTATGATATATAAATTTTGTGA 1429
Db      1002 TTTTAAATAAATAAATAANATATTTTATTTATTAATTTTAAATAATATATTAATAA 943
Qy      1430 TTTTGATTTTCATAGATTAATAATCTA--TGTTAATATGATATAATAATTTTAAATAA 1487
Db      942 AATTTATTTAATTAATAAANAATTAATAATTTAATAATATTTAATAATAAATAAATAA 883
Qy      1488 TACTATATTT--AATCTGATTAGTCGATTACCGCTTTTATATATTTTACAACTAGT 1544
Db      882 AAATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATTT 823
Qy      1545 AATATGAATAAATCAGTTATCTGAAAGCAATAATATCTTTGTGAAACAGCGTTCGGTC 1604
Db      822 ATNATNAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 763
Qy      1605 AATGGGAAGTTCATGTGATTCATAGTTTAAATAAATAAATAAATAAATAAATAAATAA 1664
Db      762 ATTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 703
Qy      1665 TTATTTTGTGTTTCAGAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAA 1704
Db      702 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 663

```

```

RESULT 7
CC262481/c 1202 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-167M9_Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9,
DEFINITION genomic survey sequence.
ACCESSION CC262481
VERSION CC262481.1 GI:30607397
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1202)
AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Kremetzki, C., Hagginsbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Nardis, E. and Wilson, R.
Gallus gallus BAC End Reads
TITLE Unpublished (2003)
JOURNAL Contact: Richard K. Wilson
COMMENT Genome Sequencing Center
Washington University School of Medicine

```

```

FEATURES
    source             1..1202
                        /organism="Gallus gallus"
                        /mol_type="genomic DNA"
                        /strain="Red Jungle Fowl"
                        /db_xref="taxon:9031"
                        /clone="CH261-167M9"
                        /sex="female"
                        /cell_line="UCD001, inbred 256"
                        /clone_lib="CH261"
                        /notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                                CH261 Female Chicken library - for library and clone
                                ordering information: http://www.chori.org/bacpac"

```

ORIGIN

```

Query Match      4.5%; Score 93; DB 8; Length 1202;
Best Local Similarity 47.3%; Pred. No. 6.6e-07;
Matches 393; Conservative 0; Mismatches 420; Indels 17; Gaps 4;

Qy      886 TCTTTAAAGTAAATGTAACCTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACTT 945
Db      1174 TATNAANAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1115
Qy      946 TTACAAGAATTTATATTAACAGAAATCATTTTATAACATGCTCTCGCGTGCATTATA 1005
Db      1114 ATAAAAATTTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1055
Qy      1006 ATAGGATCACTTACTGATCATCCATTTAAACCTTGTAAAAACAATTTCAATGAGATAA 1065
Db      1054 AAAATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1003
Qy      1066 ATATCTTCAATGAAAGAGGACAATGCTCTTTGAAAAACAATAGGTACTCCCTCC 1125
Db      1002 ATATNTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 943
Qy      1126 GTCCCTCTGAAATGTATACATATGGATTGGACAGGAGACTAAGAAAAATGTATAAAGTA 1185
Db      942 ATATTTTATTAATTAATATATAATAATAATAATAATAATAATAATAATAATAATAA 883
Qy      1186 ATGTAGAGTAAAAAGAGAGAGAGAGAGTGGGTAAAGTAGCGGGACCCCAATATAT 1245
Db      882 ATTAATAATTTATAAATAATAATAATAATAATAATAATAATAATAATAATAATAA 823
Qy      1246 AATTGATAGATTTAGAAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTATTTATATATAA 1305
Db      822 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 763
Qy      1306 AAATTTTACTATTTTGAGAAAGTTTTCGAAATGTATAGAAATTTGAGTGGGACATCCATAAAG 1365
Db      762 TAATAAATTTAAATATAAATAATAATAATAATAATAATAAATAAATAAATAAATAA 703
Qy      1366 GAAAGTGTATAGAAATTTAAATGGGACAGAGAGGAGTAATACCTTTATGATATATAAATTTT 1425
Db      702 AANAA--TAAATAATNAATTTTNAATATTTTATAATAATAATAATAATAATAAATAA 645
Qy      1426 GTTATTTTGTATTCATAGATTTATAAATCTATGTTAATAAGTAATAATAAATTTTAAAT 1485
Db      644 TATTTATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 585
Qy      1486 AATACTATATTAATTTCTGATTAGTCGATTACCGCTTTTATAAATTTTACAATACTAGTA 1545
Db      584 AAAAAAATAATTTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 531
Qy      1546 ATATGAATAAATCAGTTATCTGAAAGCAATAAATATCTTTGTAAAAACAGCGTTCGGTCA 1605
Db      530 ATATATAATAATTTTATAAATAATTTNAATAAATAAATAAATAAATAAATAAATAAATAA 471

```

```

Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.

```

Location/Qualifiers

Qy 1606 AATGGAGTTCATGTGATTCAATAGT-TTAAATATAAAGTAAATTTTAAATTAATTTG 1664
 Db 470 AAAAAATAATAATATATATTTTAAATTTTATTTAAATATAAATNAAATTAATTT 411
 Qy 1665 TTATTTTGTTCAGAAATTTAAATTAATTTGAGCATGGGAAGTTCA 1714
 Db 410 TTTTATATAAATATTAATATATATATAATAAATTAATTTATATAAATTA 361

RESULT 8
 CNS0021J 1101 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL061936
 VERSION AL061936.1 GI:4940214
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osogawa and
 Aaron Mammoss in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain V2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 Location/Qualifiers
 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR05N11"
 /clone_lib="RPCI-98"
 /note="end : TET3"

ORIGIN
 Query Match 4.5%; Score 92.4; DB 9; Length 1101;
 Best Local Similarity 39.4%; Pred. No. 8.5e-07;
 Matches 349; Conservative 65; Mismatches 469; Indels 2; Gaps 2;
 Qy 868 AGTTGTTGAGCTCATTTCTTTAAAGTATGATCTAAAGCATAGAAATT 927
 Db 138 AGTGTTCATGAAAGTTTAAATATATATATTAACGAAATCATTTTATAACATGT 197
 Qy 928 TAGTACAGGTTAAACATTTTACAGAGATTTATATTAACGAAATCATTTTATAACATGT 987
 Db 198 NAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA 257
 Qy 988 CTCTCGGCTGTCAATATAATAGGATCACCTTACTGATCATCCATTAAACCTGTTAAAA 1047
 Db 258 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA 317
 Qy 1048 CAAATTCATGATATAAATATCTTACATGAAAGAGGCAATGTCTCTTTGAAAAA 1107
 Db 318 CAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA 377

Qy 1108 CAAATAGGTACTCCCTCCGCTCTGAAATGTATACATATGGATTGGACACGAGACTA 1167
 Db 378 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA 437
 Qy 1168 AGAAAAATGTATAAGTAATGTAGAGTAAAAAGAGAAAGAAAGTGGTAAAGTAG 1227
 Db 438 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAA 497
 Qy 1228 CGGGACCCCAATATATATATATGATAGATTAGAAAAAGTAGTTGAAAGTAGTGGGT 1287
 Db 498 AAAAAAANAAAAAANAAAAAATATAATTTTATTTTAAATTAATTTTATTTTATTT 557
 Qy 1288 GGGATTTTATATATATAAATTTACTATTTTGGAGAAAGTTTGGAAATGTATAGAAATGA 1347
 Db 558 TWTATATTTTAAATTTTAAAWAAATTTAATAAANAATTTTAAATTTTAAATTTAAWAA 617
 Qy 1348 GTGGGACATCCATAAAGGAAAGTGTATAGAAATTAATGGGACAGAGGGAGTAATACCTT 1407
 Db 618 AAAAAAATTTTAAAWAAATTTTATTTTAAATTAATAAANAATTAATAAATTTTAA 677
 Qy 1408 TATGATATAT-AAAATTTTGTATTTTATTTTATTTTATATATATATATATATATG 1466
 Db 678 TWTATWAAATWAAATTTTAAAAAATTTTAAATTTTATTTTAAATTTTATTTATTTT 737
 Qy 1467 ATAATATAATTTTAAAAAATAATATACTATATTAATTTCTGATTAGTTCGATTACCGCTTTAT 1526
 Db 738 AAAAAATWATWATWATTAATTTTAAATTAATTAATAAATTAATAAATTTTAAATTTT 797
 Qy 1527 AATTTTACAATCTGAGTAAATATGAATAAATCAAGTATCTGAAAGCAATAATATCTTT 1586
 Db 798 TTTAAATTTAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 857
 Qy 1587 GTAAACACGCTTCGGTCAATGGGAGTTCATGTGTTCAATAGTTTAAATATAAAG 1646
 Db 858 TTAATTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 916
 Qy 1647 TAAATTTTAAATTTTATTTTATTTTCTGAGAAATTTAAAAATTAATTTAGACATGG 1706
 Db 917 AAATGKGAATGKKKKGKTKARKKRGKTTAAATDGTAAAGAKBATAAKKTKTKATKK 976
 Qy 1707 GAAGTTCACGGGATCATTTGAGGACGACTAGACTGTTTGAACAAAT 1751
 Db 977 TAAAAAADKAKGAAAAAATANGAARGKATGAAAAAGTGKDAATAT 1021

RESULT 9
 CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
 DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL069706
 VERSION AL069706.1 GI:4949849
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 1101)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osogawa and
 Aaron Mammoss in Pieter de Jong's laboratory in the Department of


```

Db      936 AAAAAATAATAAAAAATAAATAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 995
Qy      1382 AAATGGGACAGGAGTAAATACCTTTATGATATATAAAATTTTGGTATTTGATTTTCAT 1441
Db      996 AAAAAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 1055
Qy      1442 AAG-----ATTATAAATCTATGTTATAATGATATAATTAATTTTAAAAATAAATACTATAT 1495
Db      1056 AAGAAAAAATAAAAAATAAATAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 1115
Qy      1496 TAATCTGATAGTACCGCTTTTATAATTTTACAATCTGATGATGATGATGATGATGATGAT 1555
Db      1116 TAAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1175
Qy      1556 ATCAGTTATCTGAAAGCAATAATATCTTTGTTAAAAACAGCGTTCCGTCGAATCGGAAGT 1615
Db      1176 AATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1235
Qy      1616 TCATGTGATTTCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTTGT 1675
Db      1236 AATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1295
Qy      1676 TCAGAAATTTAAATAAA 1693
Db      1296 AATAAAAAATAAATAATA 1313

```

RESULT 11

```

CD386564      1067 bp      mRNA      linear      EST 30-MAY-2003
LOCUS      AGENCOURT_14286445 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION      CD386564
VERSION      CD386564.1 GI:31222277
KEYWORDS      EST.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1067)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgaps-remail.nih.gov

Tissue Procurement: Dr. Jamie Thompson, University of WI

cDNA Library Preparation: Gina Zastrow-Hayes

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDKM74 row: e column: 02

High quality sequence start: 8

High quality sequence stop: 462.

Location/Qualifiers

1. 1067

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic trophoblasts, made from WA01 stem

cells"

/lab_host="DH10B Tona"

/clone_lib="NIH MGC 173"

/note="vector: pDONR201; Site 1: attP2; Site 2: attP1;

LIBR-PRIMING - oligo dT; METHOD - full-length enriched;

LIBR_PROVIDER - Bradfield"

FEATURES

source

```

Best Local Similarity 44.7%; Pred. No. 3.9e-06;
Matches 305; Conservative 0; Mismatches 376; Indels 1; Gaps 1;
Qy      1015 ACTTACTGATCATCCATTAAACCTTGTAAAAACAATTCATGAGATATAAATATCTTAC 1074
Db      259 ACTCAATGAAAAACCATGATAATTCCTTTGTATATAAAATAAACATTTTGAATAAACCAAAA 318
Qy      1075 AATGAAAAAGAGGACAAATGCTCTTTGAAAAACAATAGTACTCCCTCGCTCCCTCTG 1134
Db      319 AAAAAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 378
Qy      1135 AATGTATACATATGGATTCGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGT 1194
Db      379 AAAAAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 438
Qy      1195 AAAAAAGAGAGAGAGAGAGAGTGGTAAAGTAGCGGACCCACCACCAATATATAATTTGATAG 1254
Db      439 AAAAAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 498
Qy      1255 ATTAGAAAAAGTAGTTGAAAGTAGTGGTGGGAGTATTTTATATATATAAATAATTTTACT 1314
Db      499 GGANNAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 558
Qy      1315 ATTTGAGAAAGTTTGAATGTATAGAAATGTAGATGGGACATCCATAAAGGAAAGTGTG 1374
Db      559 TTTATTTTATAAAAAATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 618
Qy      1375 TAGAATTAATGGACAGAGGAGTAACTATCTTTATGATATATAAATAATTTTGTATTTTG 1434
Db      619 NAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 677
Qy      1435 ATTTCAATAGATTTAATCTCTGTTATATATGATATATAATTTTAAAAATAAATACTATA 1494
Db      678 TTTAATATATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 737
Qy      1495 TTAATCTGATAGTCCGCTTTTATTAATTTTCACTACTGAGTAATATGAATA 1554
Db      738 TTTTNTNTATTTATATATATAAATAATAAATAAATAAATAAATAAATAAATAAATAA 797
Qy      1555 AATCAGTTATCTGAAAGCAATAATAATCTTTGTAAACAGCGTTCGGTCAATGGGAG 1614
Db      798 ATAATATATATATAAANAATAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 857
Qy      1615 TTCATGTGATTCAATAGTTTAAATAAAGTAAATTTTAAATTAATTTGTTTATTTTGT 1674
Db      858 NNTATATCATATAATATAAANAATAATTTATATAAATAAATAAATAAATAAATAAATAA 917
Qy      1675 TTCAGAAATTTAAATAAATAA 1696
Db      918 TATATTTAATAATAAATAAATA 939

```

RESULT 12

```

AG347098      1268 bp      DNA      linear      GSS 02-JUN-2004
LOCUS      Mus musculus molossinus DNA, clone:MSMg01-142I02.T7, genomic survey
DEFINITION      sequence.
ACCESSION      AG347098
VERSION      AG347098.1 GI:47920408
KEYWORDS      GSS.
SOURCE      Mus musculus molossinus
ORGANISM      Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE      Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS      BAC end Sequences of Library MSMg01
TITLE      Unpublished
JOURNAL      2 (bases 1 to 1268)
REFERENCE      Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS      Direct Submission
TITLE      Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
JOURNAL      and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

```

Query Match

4.3%; Score 88.8; DB 6; Length 1067;

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hatoriegsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENTS Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

Library

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

FEATURES

source

```
1..1268
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-142102.T7"
/sex="male"
/tissue types="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
```

ORIGIN

```
Query Match      4.3%; Score 88.8; DB 9; Length 1268;
Best Local Similarity 45.5%; Pred. No. 3.9e-06;
Matches 471; Conservative 0; Mismatches 551; Indels 14; Gaps 4;

Qy 1030 ATTAACCTTGTAAACAAATCAATGAGATAAATATCTTACAAATGAAAGAGGAC 1089
Db      |||||
Qy 1215 ATTAATAATNATAATAAAAAATATATAATAATAATAATAATAATAATAATA 1156
Db      |||||
Qy 1090 AATGCTCTTTGAAAAACAATAGGTACTCCCTCCCTCTGAAAT---GTATACAT 1146
Db      |||||
Qy 1155 AATAAATTATATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1096
Db      |||||
Qy 1147 ATGATTGGACACGGAGCTAAGAAAAATGTATAAGTAATGTAGAGTAAAGAAAGAG 1206
Db      |||||
Qy 1095 ATTAATAATTATTAATAATAATATTAATAATAAATAAATAAATAAATAAATA 1036
Db      |||||
Qy 1207 AAGAAAAGTGGGTAAGTAGCGGGACCCCAATATATTAATGATAGATTTAGAAAAGT 1266
Db      |||||
Qy 1035 AATAAATAATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 976
Db      |||||
Qy 1267 AGTTGAAGTAGTGGGTGGGATTTTATATATAAATAATTACTATTTTGAGAAAG 1326
Db      |||||
Qy 975 ATTTATAAATAATAATATTTATAAATAAATAAATAAATAAATAAATAAATAAATA 916
Db      |||||
Qy 1327 TTTTGAATGTATAGAATTGAGTGGGACATCCATAAAGGAAGTGATAGAATTAATG 1386
Db      |||||
Qy 915 ATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 856
Db      |||||
Qy 1387 GGACAGAGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGTATTTTCATAGAT 1446
Db      |||||
Qy 855 TTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 796
Db      |||||
Qy 1447 TATAAATCTATGTTATATGATATATAAATTTTAAATAAATAAATAAATAAATA 1506
Db      |||||
Qy 795 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 736
Db      |||||
Qy 1507 AGTCGATTACCGCTTTTATAAATTTTACAATCTAGTAAATGAATAAATCAAGTTATCT 1566
Db      |||||
Qy 735 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 676
Db      |||||
Qy 1567 GAAAAGCAAT---AATACTTTGTAAGAACAGCGTTCGGTCAAAATGGGAAGTTCATGTT 1623
Db      |||||
Qy 675 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 616
Db      |||||
Qy 1624 ATTCAATAGTTTTAATAAAGTAAATTTTAAATTAATTTGTTATTTTCTTCAGAAAT 1683
Db      |||||
```

```
Db      |||||
Qy 615 ATATTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 556
Db      |||||
Qy 1684 TTAATAAATAAATTTAGGATGGGAAGTTTCACGGGCATCATTTGA---GCAGCACTAGACT 1740
Db      |||||
Qy 555 ATAAATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 496
Db      |||||
Qy 1741 GTTTGAACAATGATGTCCGGTGATACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db      |||||
Qy 495 ATATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAG 436
Db      |||||
Qy 1801 GCATTCTAGAAATACATCTTTTCAAAATTTTCAACACACACAGCTTTAACTTTTCTT 1855
Db      |||||
Qy 435 AGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 376
Db      |||||
Qy 1856 CAACGGATTGGAATCCCTTTTCTAAACCTTTTAAATAAATAAATAAATAAATAAATA 1915
Db      |||||
Qy 375 AAAAAAATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 316
Db      |||||
Qy 1916 TTATCAACACCTCAACATTTAGTCTAGGCTACTATAAATAGGTGCTCTTGGTCTCTACT 1975
Db      |||||
Qy 315 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 256
Db      |||||
Qy 1976 ATCATCACATCAATCTTTACACACACACCTTTGAGCTTAATTTTCTACTTTTCTCAGCA 2035
Db      |||||
Qy 255 TAATTATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 196
Db      |||||
Qy 2036 ATAACTATTTCTAATAAT 2051
Db      |||||
Qy 195 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 180
Db      |||||

RESULT 13
CGT57503/c
LOCUS      1392 bp DNA linear GSS 24-OCT-2003
DEFINITION P052-4-C08 za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION  CGT57503
VERSION     CGT57503.1 GI:37986131
KEYWORDS   GSS.
SOURCE      Pristionchus pacificus
ORGANISM   Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE  1 (bases 1 to 1392)
AUTHORS   Stinivaasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE      An integrated physical and genetic map of the nematode Pristionchus
            pacificus
JOURNAL    Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE    22835951
PUBMED     12884007
COMMENT    Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends
            Location/Qualifiers
                1..1392
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="California"
                /db_xref="taxon:54126"
                /clone_lib="Ppa EcoRI BAC Library"
                /notes="The library was generated by a partial digest of
                the genomic DNA with EcoRI and cloning into the BAC
                vector."
```

Query Match 4.3%; Score 88.8; DB 9; Length 1392;

```
Best Local Similarity 43.8%; Pred. No. 3.9e-06;
Matches 543; Conservative 0; Mismatches 672; Indels 24; Gaps 3;

Qy 481 ATGCTCAGCCATCAAAATTCGACAAACCCGACACACACTCTATCCAGCTACTACTTTT 540
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1359 ATAATAAATATATAAATAATAATAATAATAATAATAATAATAATAATAATAATA 1300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 541 TGGCCGAATGCTCTCAAAATGTTTTTATATGTAATAATATGCCCCCAGGATAGT 600
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1299 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 601 AAAATCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATAT 660
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1239 ATTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 661 AATACTTTTACGACACAGAGACTTAGTCAAAATGAGCGCTGGTAAACAGCCTAGCT 720
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1179 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 721 TGGTCACTGATAAATAGATAATCTGTAGTATAATATAGTAGGATCTACAATGACATTA 780
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1119 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1060
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 781 ATTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTAGTAAACAGAAACAGGTA 840
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1059 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1000
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 841 AAAAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 999 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 940
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 901 TAAACTGATCTAAAGCACATAGA-----AATTTAGTACAGGTAAACCTTTTAC 949
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 939 AATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 880
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 950 AAGATTTATATTAACGAAATCATTTTATAACATGTCTCTCGGCTGCTCATATATA 1009
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 879 AATAAATATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 820
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1010 GGATCACTTCTGATCATCAATTAACCTTGTGTAAACCAAAATTCATGAGATAAATAT 1069
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 819 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 760
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1070 CTTACAAATGAAGAGGACAAATGCTCTCTTTGAAAAAACAATAGTACTCCCTCC 1129
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 759 ATTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 700
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1130 CTCTGAATGTATACATATGATTTGGACGAGACTAAGAAAAATGTATAAAGTAAT 1189
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 699 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 640
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1190 AGATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1249
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 639 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 580
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1250 GATAGATTGAGAAAGTAGTTGAAA-----GTAGTGGGTGGGTGTTTATATATA 1305
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 579 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 520
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1306 AAATTTACTATTTTGAGAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAA 1365
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 519 ATTATATATATTTATATAATATAAATAAATAAATAAATAAATAAATAAATAA 460
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1366 GAAAGTGTATAGATTAATTTGGGACAGGGAGGTAAATACCTTTATGATATATAA 1425
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 459 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 400
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1426 GTTATTTTGTATTTTCAAGATTATAAATCTATGTTTATAATGATATAATATTTT 1485
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 399 AATAATTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 340
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1486 AATACTATATAATTTCTGATTTAGTCGATTTACCGCTTTTATAATTTTACAATA 1545
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

339 AAATATATATAATTAATTTATATATTTATATATAAATATATAAATTTATAAATATATAT 280
1546 ATATGATTAATCAGTTTATCTGAAAGCAAAATATATCTTTTGTAAAAACAGGTTCCGTC 1605
279 TAAAAATTTATATATAATATATAAATAAATATATAAATTTTAAAAATAAAAAATAAATA 220
1606 AATGGGAAGTTCATGTGTATTTCAATAGTTTTTAATATAAAGTAAAAATTTTAAAAATTTGT 1665
219 AAATATATATAATTTT-----TAATAAATTTAAAAAATAAATTTAAAAATATAATA 169
1666 TATTTTGTTCAGAAATTTAAATAAATAAATTTATTTGAGCAT 1704
168 TATTTATATAAATAAATTTAAAAAATAAATAAATAAATAAATAAAT 130

RESULT 14
CC253231 1277 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-180N11.RM1.1 CH261 Gallus gallus genomic clone CH261-180N11,
DEFINITION genomic survey sequence.
ACCESSION CC253231
VERSION CC253231.1 GI:30589981
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1277)
AUTHORS Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TAGGACTCACTATAGGAGA
Class: BAC ends
High quality sequence start: 24
High quality sequence stop: 85.

FEATURES
source 1..1277
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-180N11"
/sex="female"
/call_line="UCD001, inbred 256"
/clone_lib="CH261"
/notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match 4.3%; Score 88; DB 8; Length 1277;
Best Local Similarity 43.9%; Pred. No. 5.4e-06;
Matches 475; Conservative 0; Mismatches 600; Indels 6; Gaps 3;

Qy 619 GTTTGTTTAAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACGCAA 678
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 41 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 100
Qy 679 GAGACTTAGGTCAAAAAATGGACGCTGGTAAACAGCCCTAGACTTGGTCTGCTGCTGCTGCT 738
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 101 GAACCTTTAAAAAANNCCNCAAAAAAATACATATTTTATAAAAAATATAATTTAAAAATA 160
Qy 739 TAATTTGTAGTATAATATAGTAGGACTACAATGACATTAATAATTTAGAGCTATTAATA 798
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 161 TATTTTAAACAATTTTGTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAATATT 220
```

Qy	799	GTTACTAATTAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAAACAAGAGCTTGTCGC	858
Db	221	 ATTTAAATTAAATTTATTAATTTAAATTTAAATTTATTTATATAAAATATATAAAATATAAAACAATA	280
Qy	859	TGTGTGTTTAGTTGGTGTGAGCTCATTTCTTTTAAAGTAATGTAAAACTGAATCTAAAGCAC	918
Db	281	 ATAAATTTATTAATTAATAAATAATATACATTTAAATAATTTTCAAAAATTTATAAAATTTATA	340
Qy	919	ATGAAATTTAGTACAGGTTAAAACTTTTTACAAGAAATTTATATATATAAACGAAAATCATTTT	978
Db	341	 TATPAAATTAATAATTTATTAATTAATAATAATAATAATAATAATAATAATAATAATAATA	400
Qy	979	ATAACATGTCCTCGCGTGCATTTATTAATAGGGATCACTTACTGATCATCCATTTAAAAACC	1038
Db	401	 ATTATTAATAAATAAATAATAATAATAATAATAATTTTATTTTAAATAAAAAAAAAATTAATAATA	460
Qy	1039	TTGTTAAAAACAAATTCATGATGAGATABAATAATCTTTACAAATGNAAGAAGGCAATGCTCT	1098
Db	461	 TTTATTTATATTTTATTTATTAATAATTTATAAAATTTATTTAAATTTAAATTTTAAATAATAATAA	520
Qy	1099	TTGAAAAACAAATPAGGTACTCCCTCCGTCCTCGAAATGTATACATATGGATTTGGACA	1158
Db	521	 ATATTAACTATATATTTTAAATAATTTAAAAATATATATTAATAATAAAATATATTATATA	580
Qy	1159	CGGA-GACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAGAGAAAGAAAGTG	1217
Db	581	 TAAATTTATTTAAATAAATAATTTAAATTTAAAAATATAAATATAAAAAATTAATAAAAAATAT	640
Qy	1218	GGTAAAGTAGCGGACCCCAATATATAATTCGATAGATTTTAGNAAAGTAGTTGCAAGTA	1277
Db	641	 ATATAATAAATAAATTAATAATTTAAATAATTAATAAAAAATAAATTTAAATAAATTTATATAAA	700
Qy	1278	GTGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTGAGAAAGTTTTCGAAATGT	1337
Db	701	 ATTAATTTTAAATTTTATATAATAATTAATTTATAAATAATTTTAAAAATTTTATAATTAAT	760
Qy	1338	ATGAAATGAGTGGGACATCCATAAAAGGAAAGTGATAGAAATTAATGGACAGAGCGGA	1397
Db	761	 AATATATATAATTTTAAATAAATTTAAATATAAATAAATAAATAATTAATTAATTTAAATAAAT	820
Qy	1398	GTAATACCTTTATGATATATAAATTTTGTGTTATTTGATTTTCATAAGATTTATAAATCTAT	1457
Db	821	 AAAATAATAAANAATATAAATAATTTTATTTATAAAAAAATAAATAAATAAATAAATAAAT	880
Qy	1458	GTTATAATGATAATATAATTTTAAAAATAACTATATATAATTAATTTAGTTCGATTAACC	1517
Db	881	 TATAAATAATACATAAATAATAATAAATTTAAATTTATATTAATTTAAATTTAAATAATAATAA	940
Qy	1518	GCCCTTTTATAATTTTACAATACTGAGTGAATATGAATAAATCAGTTTCTCGAAAGCAAAT	1577
Db	941	 ATATTTTTTATA--TAAATAAATAATATAATTTTATATAAATAAAAAAATAAATAATAAAT	997
Qy	1578	AATATCTTTGFAAACACAGCGTTCCGTCCAATGGGAAGTTCATGTGTGTAATCCAATAGTTT	1637
Db	998	 AAATTTTATATAAAAAAATAAAT--ATAAATAATTAATAATTTTATAAATAATTTCTATATTAATA	1055
Qy	1638	ATATAAAGTAAATTTTAAAAATTAATTCGTTATTTTGTTCAGAAATTTAAAAATAAATTTAT	1697
Db	1056	 ATAATATATATAAATAATTTAATTTATATATATATTAATAATAATAATAATAATAATAATAA	1115
Qy	1698	T 1698	
Db	1116	T 1116	

RESULT 15	
AG346840/c	
LOCUS	AG346840
DEFINITION	1696 bp DNA linear GSS 02-JUN-2004
	Mus musculus molossinus DNA, clone:MSMg01-142C12.T7, genomic survey
	sequence.
ACCESSION	AG346840
VERSION	AG346840.1
KEYWORDS	GI:47920150
	GSS.

```

SOURCE
ORGANISM
Mus musculus molossinus
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1696)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Shuhiro-Chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp_gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp)
Teukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACE3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
Location/Qualifiers
1..1696
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-142C12.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

```

ORIGIN

Query Match	4.3%; Score 88; DB 9; Length 1696;
Best Local Similarity	47.3%; Pred. No. 5.4e-06;
Matches	464; Conservative 0; Mismatches 501; Indels 16; Gaps 6;
Qy	715 TAGACTGGGCACCTGATATAATAGATAATCTGTAGTATAATATAGTAGGATCTACAATGCAC 774
Db	1290 TATTAATTAATAATATCTTAATAATAATAATAATAATAATAACATATATATAAAAA 1231
Qy	775 ATTAAAAATTAGAGCTATTAAATTAAGTTACTAAATAATAAGAGAGGTTAGTAAACAGAAAG 834
Db	1230 ATAAAAATAAACAAAAACAACATATATACAAATAATAATAATAATAATAAACAAAAAAA 1171
Qy	835 CAGTAAAAACAAGAGCTTCGCTGCTGTGCTGTAGTTGTTGTGAGCTCATTTCTTTAAAA 894
Db	1170 AATATAATATAAAAAATAAATAAATCTTAAACAAAAACAATAAATAACATATACATATA 1111
Qy	895 GTAATGTAAACTGATCTTAAAGCACAAGAAAAATTTAGTACAGGTTAAAACTTTTCAAGAA 954
Db	1110 ATATAAATAATCAATACATCAAAATTATAAAAAAAAATAATAAACATATA -AATAA 1052
Qy	955 TTTATATTAACGAAATCATTTTATTAACATGCTCTCGGCTGTCATTTAATAAGGATC 1014
Db	1051 ATCATATTAAAAATAATATCTATCAATAATAAACNTACAAACATAATAATAATAATCACC 992
Qy	1015 ACTTACTGATCATCCATTAACAACTCTTTAAAAACAAATTCAAATGAGATAAAAAATCTCTTAC 1074
Db	991 ATAAATAATAAAAAATAAAAAATAATAATAAATAATCTTACTAAAAATAAAAAATAAT - 934
Qy	1075 AATGAAAGAGGCAATGTCTCTTTGAAAAACAATAAGGTACTCCCTCGCTCCCTCTG 1134
Db	933 -----AATAAAAACAATATATATATATATATATATATATAAAAATAACAAATAACAATATATA 880

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2005, 13:48:55 ; Search time 355 Seconds
(without alignments)
9458.140 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtgcctacagcaca.....gcaataacattcctaataatc 2052

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241.2	11.8	2042	2	US-08-911-434A-2
2	196.6	9.6	247	2	US-08-911-434A-1
3	195.6	9.5	246	3	US-09-352-608-2
4	119	5.8	140	3	US-09-352-608-6
5	119	5.8	140	3	US-09-352-608-7
6	81.4	4.0	187169	4	US-09-949-016-12776
7	81.4	4.0	191569	4	US-09-949-016-15940
8	72.2	3.5	205044	4	US-09-949-016-15851
9	72.2	3.5	205044	4	US-09-949-016-15852
10	72.2	3.5	205044	4	US-09-949-016-15853
11	72.2	3.5	223471	4	US-09-949-016-12387
12	72.2	3.5	223471	4	US-09-949-016-12724
13	72.2	3.5	223471	4	US-09-949-016-12725
14	69.4	3.4	7218	1	US-08-232-463-14
15	68.6	3.3	1141	4	US-09-806-708B-22
16	68.6	3.3	18773	4	US-09-949-016-14164
17	68.2	3.3	1141	4	US-09-806-708B-22
18	67.8	3.3	205044	4	US-09-949-016-15851
19	67.8	3.3	205044	4	US-09-949-016-15852
20	67.8	3.3	205044	4	US-09-949-016-15853
21	67.8	3.3	223471	4	US-09-949-016-12387
22	67.8	3.3	223471	4	US-09-949-016-12724
23	67.8	3.3	223471	4	US-09-949-016-12725
24	67.2	3.3	601	4	US-09-949-016-30530
25	67.2	3.3	601	4	US-09-949-016-30531
26	67.2	3.3	601	4	US-09-949-016-37149
27	67.2	3.3	601	4	US-09-949-016-37150

Sequence 37163, A
Sequence 37164, A
Sequence 145867, A
Sequence 145868, A
Sequence 146135, A
Sequence 146136, A
Sequence 146403, A
Sequence 146404, A
Sequence 30531, A
Sequence 37150, A
Sequence 37164, A
Sequence 145868, A
Sequence 146136, A
Sequence 146404, A
Sequence 156535, A
Sequence 16110, A
Sequence 19334, A
Sequence 13248, A

ALIGNMENTS

RESULT 1

US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satoshi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION/DOCKET NUMBER: 21,066
; REFERENCE/SEQUENCE NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..2042
US-08-911-434A-2

Query Match 11.8%; Score 241.2; DB 2; Length 2042;
Best Local Similarity 72.8%; Pred. No. 5.4e-42;

Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;

Qy 1554 AAATCAGTATCTGAAAGCAAAATATATCTTTGTAAACAGCG- - - - -TTCCGGTCAATG 1609
 Db 1543 AAATTTATCTGAATGAATAACATCTTTGTAAACAAACCTGGCCAAATAGGACCATAA 1602
 Qy 1610 GGAAGTTCATGTGATTCATAGTTTAAATATAAAAGTAAATTTAAATTAATTTGTTATT 1669
 Db 1603 CCAAGTTTCAGTGTATTTCTAAATGTTAAATCTAACATGAGTATTTTCTT- -TTCAAGGT 1660
 Qy 1670 TTTGTTTTCAGAAATTTAAATAAATATTCAGCATGGAAAGTTTCACGGCATCATTCGAGC 1729
 Db 1661 ATAAGTTAAATCTTCATCAATTAACCTTTAAATTTGGACATTAATGAGCAACTTTATGCC 1720
 Qy 1730 AGCACTAGACTGTTTGAACAAATGATGTCGGGTGTACATCTATGACCTTTTCAACTCAAAC 1789
 Db 1721 CACGTTGATGTTTAAACAAACGTTTGTCCGGGTATATATTTATGACCTTTTCAACTCAAGC 1780
 Qy 1790 TAGTGAT- AATGCAATCTAGNATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
 Db 1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
 Qy 1849 TTTCTTTTCAACGGATTCGAATCCTTTTCTAAACCTTTTAAATATAAAATGCAATTTATT 1908
 Db 1841 TTTCTTTTCAAGATTAGAAATGCTTTCTTAACTTTTAAATTT- -AAAAATACATTAAT 1898
 Qy 1909 GTAATATTATCAACACTCAACATTCATGTTAGCGTACTATATAATAGGTGCTCTTGGTG 1968
 Db 1899 ATAATATTATCAACACTCAACATTCATGTTAGCGTACTATATAATAGGTGCTCTTGGTG 1958
 Qy 1969 CTCTACTATCATCATCAATCTTTACACCAAAACCTTGAGCTTAATTTTCTACTTATT 2028
 Db 1959 CTCTACTATCATCATCAATCTTTCCAGCAAAACCTTGAGCTTAATTTTCTACTTAAAT 2018
 Qy 2029 CTAGCAATAACATCTTAAATATC 2052
 Db 2019 TTTAGCAAAACATCTTAAAGGTC 2042

RESULT 2

US-08-911-434A-1
 ; Sequence 1, Application US/08911434A
 ; Patent No. 5959176
 ; GENERAL INFORMATION:
 ; APPLICANT: TORIKAI, Satomi
 ; APPLICANT: OEDA, Kenji
 ; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
 ; STREET: P.O. BOX 747
 ; CITY: FALLS CHURCH
 ; STATE: VIRGINIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 22040
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/911.434A
 ; FILING DATE: 12-AUG-1997
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stewart, Raymond C.
 ; REGISTRATION NUMBER: 21,066
 ; REFERENCE/DOCKET NUMBER: 2185-0199P
 ; TELEPHONE: (703)205-8000
 ; TELEFAX: (703)205-8050
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 247 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Daucus carota L.
 INDIVIDUAL ISOLATE: Kuroda Gosun
 FEATURE:
 NAME/KEY: promoter
 LOCATION: 1..247
 US-08-911-434A-1

Query Match 9.6%; Score 196.6; DB 2; Length 247;
 Best Local Similarity 89.6%; Pred. No. 1e-32;
 Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy 1804 TTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1863
 Db 1 TTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCTAACTTTTCTTTTAAACAGAT 60
 Qy 1864 TCGAATCCTTTTCTAAACCTTTTAAATATAAAATGCAATTTGTAATATTATCAAC 1923
 Db 61 TAGAATCGTTTCTTAACTTTTAAATTT- -AAAAATACATTAATTAATTTATCAAC 118
 Qy 1924 ACCTCAACATTTAGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCAC 1983
 Db 119 ACCTCAACATTTAGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCAC 178
 Qy 1984 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATAACATT 2043
 Db 179 ATCAATCTTCCAGCAAAACCTTGAGCTTAATCTTTCTACTTAATTTTTCAGCAAAACATT 238
 Qy 2044 CTAATATATC 2052
 Db 239 CTAAGGTC 247

RESULT 3

US-09-352-608-2
 ; Sequence 2, Application US/09352608
 ; Patent No. 6218598
 ; GENERAL INFORMATION:
 ; APPLICANT: ISHIGE, Fumiharu
 ; APPLICANT: NISHIKAWA, Satomi
 ; APPLICANT: OEDA, Kenji
 ; TITLE OF INVENTION: Plant Promoter
 ; FILE REFERENCE: 2185-0353P
 ; CURRENT APPLICATION NUMBER: US/09/352,608
 ; CURRENT FILING DATE: 1999-07-13
 ; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
 ; EARLIER FILING DATE: 1998-07-15
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 246
 ; TYPE: DNA
 ; ORGANISM: Daucus carota
 ; US-09-352-608-2

Query Match 9.5%; Score 195.6; DB 3; Length 246;
 Best Local Similarity 89.5%; Pred. No. 1.7e-32;
 Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy 1805 TCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGATT 1864
 Db 1 TCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACAGATT 60
 Qy 1865 GGAATCCTTTTCTAAACCTTTTAAATATAAAATGCAATTTGTAATATTATCAACA 1924
 Db 61 AGAATCGTTTCTTAAACCTTTTAAATTT- -AAAAATACATTAATTAATTTTATCAACA 118

Qy 1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 1984
Db 119 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCAC 178
Qy 1985 TCAATCTTACACCAACACCTTGAGCTTAATTTTCTTACTATTATCTCAGCAATTAACATTC 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATTTTCTTACTAAATTTTGGCAAAACATTC 238
Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 4
US-09-352-608-6
; Sequence 6, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-352-608-6

Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 2.8e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1918 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 1977
Db 5 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 64
Qy 1978 CATCATCATCAATCTTACACCAACCTTGAGCTTAATTTTCTTACTTATCTCAGCAAT 2037
Db 65 CATCATCATCAATCTTCCAGCACAAACCTTGAGCTTAATTTTCTTACTAAATTTTGTAGCAAA 124
Qy 2038 AACATTTCTAAATATC 2052
Db 125 AACATTTCTAAAGGTC 139

RESULT 5
US-09-352-608-7/c
; Sequence 7, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 140
; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-352-608-7

Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 2.8e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1918 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 1977
Db 140 ATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTAT 81
Qy 1978 CATCATCATCAATCTTACACCAACCTTGAGCTTAATTTTCTTACTTATCTCAGCAAT 2037
Db 80 CATCATCATCAATCTTCCAGCACAAACCTTGAGCTTAATTTTCTTACTAANTTTTGTAGCAAA 21
Qy 2038 AACATTTCTAAATATC 2052
Db 20 AACATTTCTAAAGGTC 6

RESULT 6

US-09-949-016-12776
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(187169)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

Query Match 4.0%; Score 81.4; DB 4; Length 187169;
Best Local Similarity 45.3%; Pred. No. 1.4e-07;
Matches 417; Conservative 0; Mismatches 496; Indels 8; Gaps 3;
Qy 883 ATTTCTTTAAAGTAATGTAAACTGATCTAAAGCACATAGAAATTTAGTCAGGTGCTTAA 942
Db 465 ATATATATATAAAATATATATATATATAAATATATAAATATATAAATATATAAATATAT 524
Qy 943 CTTTTCAAGAATTTATTTAAACGAAATCATTTTATACATGCTCTCTCGGCTGTCATT 1002
Db 525 ATAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 584
Qy 1003 ATAATAGGATCATCTTACTGATCATCCATTAACCTTGTGTTAAACAAATTCATGAT 1062
Db 585 TAAATATATATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 644
Qy 1063 AAAATATCTTACAAATGAAAGAGGCAATGTCTCTTTGAAAAACAAATAGGTACTCCC 1122
Db 645 ATAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATA 704
Qy 1123 TCGTCCCTCTGAAATGTATACATATGATTTGGACACGGAGACTAAGAAAAATGTATAA 1182

US-09-949-016-15851
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15851
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

Query Match 3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;
Qy 1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATCTAGAGT 1194
Db 20520 ATATATTTATATATATTTTAAATATAAATATATAAATAATATATTTATATATTTAAAT 200579
Qy 1195 AAAAAAGAAAGAGAAAGTGGTAAAGTAGCGGACCCCAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAAAATATATTTTATATATATTTTAAATATAAATAATATATTTAT 200639
Qy 1251 ATAGATTAGAAAAGTAGTTGAAAGTAGTGGTGGGAGTTTATATATAAATAATTT 1310
Db 200640 ATATATTTTAAATATAAATAATATAAATAATATATTTTAAATATAAATAATATA 200699
Qy 1311 TACTATTTTCAGAAAGTTTGAATGTATAGATTGAGTGGACATCCATAAAGGAAAG 1370
Db 200700 AAATATATTTATATATATTTTAAATATAAATAATATAAATAATATATTTTAAAT 200759
Qy 1371 TGTATAGAATTAATGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATAAATAATATATTTATATATATTTTAAATATAAATAATATAAATA 200819
Qy 1431 TTTGATTTCAAGATTATAAATCTATGTTATATATGATAATATAAATTTTAAAAATAATAC 1490
Db 200820 TTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAATATAAATA 200879
Qy 1491 TATATTAATTTCTGATTAGTCGATTACCGCTTTTATATATTTTAAATATAAATAATATA 1550
Db 200880 TATTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAATATA 200939
Qy 1551 AATAAATACGTTATCTGAAAGCAAAATAATCTTTGTAAAAACAGCGCTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAAATA 200999
Qy 1607 -----ATGGGAAGTTCAATGTTTCAATAGTTTAAATATAAAGTAATAATTTTAAAT 1661
Db 201000 TATATATTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAAATA 201059
Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATTAATTTGA 1700
Db 201060 TATATATATTTTAAATATAAATAATATAAATAATATATTTTA 201098

RESULT 9

US-09-949-016-15852
; Sequence 15852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15852
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15852

Query Match 3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;
Qy 1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATCTAGAGT 1194
Db 20520 ATATATTTATATATATTTTAAATATAAATAATATAAATAATATATTTTAAAT 200579
Qy 1195 AAAAAAGAAAGAGAAAGTGGTAAAGTAGCGGACCCCAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAAAATATATTTTATATATATTTTAAATATAAATAATATAAATA 200639
Qy 1251 ATAGATTAGAAAAGTAGTTGAAAGTAGTGGTGGGAGTTTATATATAAATAATTT 1310
Db 200640 ATATATTTTAAATATAAATAATATAAATAATATATTTTAAATATAAATAATATA 200699
Qy 1311 TACTATTTTCAGAAAGTTTGAATGTATAGATTGAGTGGACATCCATAAAGGAAAG 1370
Db 200700 AAATATATTTATATATATTTTAAATATAAATAATATAAATAATATATTTTAAAT 200759
Qy 1371 TGTATAGAATTAATGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATAAATAATATATTTATATATATTTTAAATATAAATAATATAAATA 200819
Qy 1431 TTTGATTTCAAGATTATAAATCTATGTTATATATGATAATATAAATTTTAAAAATAATAC 1490
Db 200820 TTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAATATAAATA 200879
Qy 1491 TATATTAATTTCTGATTAGTCGATTACCGCTTTTATATATTTTAAATATAAATAATATA 1550
Db 200880 TATTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAATATA 200939
Qy 1551 AATAAATACGTTATCTGAAAGCAAAATAATCTTTGTAAAAACAGCGCTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAAATA 200999
Qy 1607 -----ATGGGAAGTTCAATGTTTCAATAGTTTAAATATAAAGTAATAATTTTAAAT 1661
Db 201000 TATATATTTTAAATATAAATAATATAAATAATATAAATTTTAAATATAAATAAATA 201059
Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATTAATTTGA 1700
Db 201060 TATATATATTTTAAATATAAATAATATAAATAATATATTTTA 201098

RESULT 10

```
US-09-949-016-15853
; Sequence 15853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15853
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15853
```

```
Query Match      3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

QY 1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTACAGT 1194
Db 200520 ATATATTTATATATATTTTAAATATAAATATATAAATATATTTATATATATTTAAAT 200579

QY 1195 AAAAAAGAGAGAGAAAAAGTGGGTAAAGTAGCGGACCCACCATAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAAAATATATTTATATATATTTTAAATATAAATATATAAATAATATTTAT 200639

QY 1251 ATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAATAAT 1310
Db 200640 ATATATTTTAAATATAAATATATAAATATATATTTTATATATATTTTAAATATAAATATATA 200699

QY 1311 TACTATTTTCGAGAAAGTTTGAATGTATAGAATTGAGTGGGACATCCATAAAGGAAAG 1370
Db 200700 AATATATATTTATATATATTTTAAATATAAATATATAAATATATTTATATATATTTTAA 200759

QY 1371 TGTATAGAATTTAAATGGGACAGGGAGTAATACCTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATATAAATATATTTATATATATTTTAAATATAAATATATAAATAATATATA 200819

QY 1431 TTGATTTTCATAGATTTATAAATCTATGTTATAATGATATAATATTTTAAATAATATAC 1490
Db 200820 TTTTAAATATAAATATATAAATAATATATATATTTTAAATATAAATATATAAATAATATA 200879

QY 1491 TATATTAATTTCTGATTTAGTCGATTACCGCTTTTATATAATTTTACAATACCTGAGTAATG 1550
Db 200880 TATTTTAAATATAAATATATAAATAATATATATATTTTAAATATAAATATATAAATAATATA 200939

QY 1551 AATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTGTAACACGCGTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATA 200999

QY 1607 -----ATGGGAGTTCATGTTATCAATAGTTTAAATATAAAGTAAATTTTAAATTA 1661
Db 201000 TATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAA 201059

QY 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATTTATGA 1700
Db 201060 TATATATATTTTAAATATAAATAATATAAATAATATTTTA 201098
```

RESULT 11

```
US-09-949-016-12387
; Sequence 12387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12387
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12387
```

```
Query Match      3.5%; Score 72.2; DB 4; Length 223471;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

QY 1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTACAGT 1194
Db 170947 ATATATTTATATATATTTTAAATATAAATATATAAATATATTTATATATATTTTAAAT 171006

QY 1195 AAAAAAGAGAGAGAAAAAGTGGGTAAAGTAGCGGACCCACCATAATATATAA----TTG 1250
Db 171007 ATAAATATATAAAAAATATATTTATATATATTTTAAATATAAATATATAAATAATATTTAT 171066

QY 1251 ATAGATTTAGAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAATAAT 1310
Db 171067 ATATATTTTAAATATAAATATATAAATATATATTTTATATATATTTTAAATATAAATATATA 171126

QY 1311 TACTATTTTCGAGAAAGTTTGAATGTATAGAATTGAGTGGGACATCCATAAAGGAAAG 1370
Db 171127 AATATATATTTATATATATTTTAAATATAAATATATAAATATATTTATATATATTTTAA 171186

QY 1371 TGTATAGAATTTAAATGGGACAGGGAGTAATACCTTTATGATATATAAATTTTGTAT 1430
Db 171187 TATAAATATATAAATAATATATTTATATATATTTTAAATATAAATATATAAATAATATATA 171246

QY 1431 TTGATTTTCATAGATTTATAAATCTATGTTATAATGATATAATATTTTAAATAATATAC 1490
Db 171247 TTTTAAATATAAATATATAAATAATATATATATTTTAAATATAAATATATAAATAATATA 171306

QY 1491 TATATTAATTTCTGATTTAGTCGATTACCGCTTTTATATAATTTTACAATACCTGAGTAATG 1550
Db 171307 TATTTTAAATATAAATATATAAATAATATATATATTTTAAATATAAATATATAAATAATATA 171366

QY 1551 AATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTGTAACACGCGTTCGGTCAA---- 1606
Db 171367 TATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATA 171426

QY 1607 -----ATGGGAGTTCATGTTATCAATAGTTTAAATATAAAGTAAATTTTAAATTA 1661
Db 171427 TATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATA 171486

QY 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATTTATGA 1700
```


Db 171487 TATATATATTTAAATATAAAATATATAAAATATATTTA 171525

RESULT 12

US-09-949-016-12724
; Sequence 12724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12724
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724

Query Match 3.5%; Score 72.2; DB 4; Length 223471;

Best Local Similarity 48.0%; Pred. No. 1.3e-05; Mismatches 288; Indels 13; Gaps 2;

Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

Qy 1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAAATGTAGAGT 1194

Db 170947 ATATATTTATATATATTTTAAATATAAATATAAATAATATTTATATATATTTTAAAT 171006

Qy 1195 AAAAAGAAAGAAAGTGGTAAAGTAGCGGGAGCCCAACCAATATAAA----TTG 1250

Db 171007 ATAAATATATAAAATATATTTTATATATATTTTAAATATAAATATAAATAATATTTAT 171066

Qy 1251 ATAGATTTAGAAAAGTAGTTCGAAAGTAGTGGTGGGAGTATTTTATATATAAATAAT 1310

Db 171067 ATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 171126

Qy 1311 TACTATTTTGAGAAAGTTTGGAAATGTATAGAAATGGAGTGGGACATCCATAAAGGAAAG 1370

Db 171127 AAATATATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATA 171186

Qy 1371 TGTATAGAAATTAAGTGGACAGAGGAGTAATACCTTTATGATATATAAAATTTTGTAT 1430

Db 171187 TATAAATATAAATAATATATTTTATATATATTTTAAATATAAATATAAATAATATAAATA 171246

Qy 1431 TTTCGATTTTCAAGATTAATACTCTGTTATATGATAAATAATTTTAAATAAATAATAC 1490

Db 171247 TTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAAATA 171306

Qy 1491 TATATTAATCTGATTCGATTACCGCTTTTATATATTTTAACTACTGAGTAATATG 1550

Db 171307 TATTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAAATA 171366

Qy 1551 AATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAACACGGTTCGGTCAA---- 1606

Db 171367 TATATTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAAATA 171426

Qy 1607 -----ATCGGAGTTCATGCTGTTCAATAGTTTTAAATATAAAGTAAATTTTAAATTA 1661

Db 171427 TATATATTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAA 171486

Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATAAATAAATTTTGA 1700

Db 171487 TATATATATTTAAATATAAAATATATAAAATATATTTA 171525

RESULT 13

US-09-949-016-12725
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12725
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725

Query Match 3.5%; Score 72.2; DB 4; Length 223471;

Best Local Similarity 48.0%; Pred. No. 1.3e-05; Mismatches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

Qy 1135 AAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAAATGTAGAGT 1194

Db 170947 ATATATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATA 171006

Qy 1195 AAAAAGAAAGAAAGTGGTAAAGTAGCGGGAGCCCAACCAATATAAA----TTG 1250

Db 171007 ATAAATATATAAAATATATTTTATATATATTTTAAATATAAATATAAATAATATTTAT 171066

Qy 1251 ATAGATTTAGAAAAGTAGTTCGAAAGTAGTGGTGGGAGTATTTTATATATAAATAAT 1310

Db 171067 ATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATA 171126

Qy 1311 TACTATTTTGAGAAAGTTTGGAAATGTATAGAAATGGAGTGGGACATCCATAAAGGAAAG 1370

Db 171127 AAATATATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATA 171186

Qy 1371 TGTATAGAAATTAAGTGGACAGAGGAGTAATACCTTTATGATATATAAAATTTTGTAT 1430

Db 171187 TATAAATATAAATAATATATTTTATATATATTTTAAATATAAATATAAATAAATAAATA 171246

Qy 1431 TTTCGATTTTCAAGATTAATACTCTGTTATATGATAAATAATTTTAAATAAATAATAC 1490

Db 171247 TTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAAATA 171306

Qy 1491 TATATTAATCTGATTCGATTACCGCTTTTATATATTTTAACTACTGAGTAATATG 1550

Db 171307 TATTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAAATA 171366

Qy 1551 AATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAACACGGTTCGGTCAA---- 1606

Db 171367 TATATTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAAATA 171426

Qy 1607 -----ATCGGAGTTCATGCTGTTCAATAGTTTTAAATATAAAGTAAATTTTAAATTA 1661

Db 171427 TATATATTTTAAATATAAATAATATAAATAATATAAATATAAATATAAATAAATAAATAA 171486

Qy 1662 TTGTTATTTTGGTTTCAGAAATTTAAATTAATTAATTAATGA 1700
Db 171487 TATATATATTTTAAATAATAATATATAAAATATATTTA 171525

RESULT 14

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ5pt-Fls
US-08-232-463-14

Query Match 3.4%; Score 69.4; DB 1; Length 7218;
Best Local Similarity 4.6%; Pred. No. 2.4e-05;
Matches 16; Conservative 212; Mismatches 123; Indels 0; Gaps 0;

Qy 1051 ATTCAATGAGATAAAATATCTTACAATGAAAGAGACAATGTCCTTTGAAAAACAA 1110
Db 1461 AGTTAAGAGATAGAAGAAATTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1402
Qy 1111 ATAGGTACTCCCTCGCTCCTGCTGAATGATATACATATGAGTTGGACGAGACTAAGA 1170
Db 1401 RRR 1342
Qy 1171 AAAATGATAAAGTAAGTAGAGTAAAGAAAGAGAAAGAAAGTGGTAAAGTACGG 1230
Db 1341 RRR 1282
Qy 1231 GACCACCAATATATAATATGATAGTTAGAAAAAGTAGTTGAAAGTAGTGGTGGG 1290
Db 1281 RRR 1222

Qy 1291 ATTTTATATATTATAAAATTTACTATTTTCAGAAAGTTTGAATGTATAGAAATTGAGTG 1350
Db 1221 RRR 1162
Qy 1351 GGACATCCATAAAAGGAAGTGTATAGAATTAATGGGACAGAGGAGTAA 1401
Db 1161 RRR 1111

RESULT 15

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match 3.3%; Score 68.6; DB 4; Length 1141;
Best Local Similarity 13.1%; Pred. No. 2.4e-05;
Matches 141; Conservative 392; Mismatches 536; Indels 10; Gaps 5;

Qy 538 TTTTGGCCGAATGCTTCTCAAAATGTTTATATATGTTAAATATATGCCCCATCCAGATA 597
Db 1129 TTKTKTYKANNNNNNNGKDNRMDATKWSATGTAWTTHAKRGATMCWYWTGTN 1070
Qy 598 AGTAAAAATCCCGTTTAAACAGTTTGTATATATATATATGTTTACATTTCAAGAGATATT 657
Db 1069 RRMCMRTYAMRTWYTRSNANWSCATKBMWMTKMYATKYRTA-WYANWCAWRNNMMCAT 1011
Qy 658 CGTAATACTTTTAGACGACAGAGACTTAGTCAAAAATCGACGCTGGTAAACAGCCTAG 717
Db 1010 NGYAKSCATNNAMWYATTRWAAYAAAKWARWAGNNMRMYGAAAGNKWGCMAAMATMGBW 951
Qy 718 ACTTGGTCACTGATAAATAGATAAATTTGTTAGTATATATATAGTAGGATCTCAATGACATT 777
Db 950 ADTAGKMCNNNNNNWTTDVRMMAM--KAKNNNNNNAYWTACYNRAATNNKMATHWMKWTH 893
Qy 778 AAAATTAGAGCTATTAAATTAAGTTACTATAATAAGAGAGGTTAGTAAACAGAAAGCAG 837
Db 892 GAHSKRTRHRTTCRRRTKYNNNNNNNARTVYVYHHAARRMWNWTTTNNNNNNNNNNAC 833
Qy 838 GTAAAAACAAGAGCTTCTGCTGCTGTTAGTTGTTGAGAGCTCATTTCTTTTAAAAAGTA 897
Db 832 RNTRTW--WABWKHSWCNN 775
Qy 898 ATGTAACACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACACTTTTACAAGAATTT 957
Db 774 NNYMHAAVTTTHTDWCYKTYMMNTWYDMWMTTMBTTTTNNMTTSTNNMTNNNNNNWACTNN 715
Qy 958 ATATTAAACCAAAATCATTTTATAACATGTCCTCGCTGCTCATTTATAATAGGATCACT 1017
Db 714 NNNMMKAYTAHANWGCWNNNTDARTNNNTVMRRRWNTYKTRVYSTTR--HHYT 658
Qy 1018 TACTGATCATCCATTAAACCTTTGTTAAAAACAAATTCAAATGAGATAAAATATCTTACAAT 1077
Db 657 GATNN 598

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2005, 16:47:36 ; Search time 1690 Seconds
(without alignments)
10028.076 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtggcctacagcaca.....gcaatacattctaaatc 2052

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/PCT NEW PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT NEW PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US10_NEW PUB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US10_NEW PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US11_NEW PUB.seq.*
27: /cgn2_6/ptodata/2/pubpna/US60_NEW PUB.seq.*
28: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99.4	4.8	8056	21	US-10-473-126-386
C 2	93	4.5	3673778	17	US-10-312-841-1
3	91.8	4.5	8056	21	US-10-473-126-240
C 4	91.6	4.5	8056	21	US-10-473-126-386
C 5	88.6	4.3	8056	21	US-10-473-126-240

6	83	4.0	158001	18	US-10-211-179-11	GENERAL INFORMATION
7	79.4	3.9	74665	21	US-10-719-993-6854	Sequence 6854, App
C 8	75	3.7	11745	16	US-10-240-453-206	Sequence 206, Appl
C 9	71.8	3.5	176096	24	US-10-981-277-34	Sequence 34, Appl
C 10	71.6	3.5	593	13	US-09-925-065A-323654	Sequence 323654,
C 11	71.2	3.5	8222	22	US-10-486-319A-65	Sequence 65, Appl
C 12	71	3.5	5930	16	US-10-311-455-490	Sequence 490, Appl
C 13	70.5	3.4	3673778	17	US-10-312-841-1	Sequence 1, Appl
C 14	70.2	3.4	38678	24	US-10-893-315-136	Sequence 136, App
C 15	70.2	3.4	38684	24	US-10-893-315-154	Sequence 13786,
C 16	70	3.4	3252	14	US-10-027-632-113786	Sequence 113787,
C 17	70	3.4	3252	14	US-10-027-632-113787	Sequence 113788,
C 18	70	3.4	3252	14	US-10-027-632-113788	Sequence 113787,
C 19	70	3.4	3252	14	US-10-027-632-113786	Sequence 113786,
C 20	70	3.4	3252	18	US-10-027-632-113787	Sequence 113787,
C 21	70	3.4	3252	18	US-10-027-632-113788	Sequence 113788,
C 22	69.8	3.4	6352	18	US-10-221-613-195	Sequence 195, App
C 23	69.6	3.4	8222	22	US-10-486-319A-43	Sequence 43, Appl
C 24	69.4	3.4	16258	18	US-10-257-166-120	Sequence 120, App
C 25	69	3.4	3673778	17	US-10-312-841-2	Sequence 2, Appl
C 26	68.6	3.3	536	24	US-10-893-315-1193	Sequence 1193, Ap
C 27	68.6	3.3	546	24	US-10-893-315-1192	Sequence 1191, Ap
C 28	68.6	3.3	38678	24	US-10-893-315-136	Sequence 136, App
C 29	68.6	3.3	38684	24	US-10-893-315-154	Sequence 154, App
C 30	68.6	3.3	38684	24	US-10-893-315-154	Sequence 154, App
C 31	68.6	3.3	49979	20	US-10-741-601-5746	Sequence 5746, Ap
C 32	68.6	3.3	49979	22	US-10-741-600-17905	Sequence 17905, A
C 33	68.2	3.3	15548	16	US-10-311-455-2128	Sequence 2128, Ap
C 34	68	3.3	11422	16	US-10-311-455-191	Sequence 191, App
C 35	68	3.3	11422	18	US-10-257-166-17	Sequence 17, Appl
C 36	67.8	3.3	15732	15	US-10-239-676-95	Sequence 95, Appl
C 37	67.8	3.3	15732	16	US-10-240-453-107	Sequence 107, App
C 38	67.6	3.3	5984	20	US-10-433-793-24	Sequence 24, Appl
C 39	67.6	3.3	6050	16	US-10-311-455-1984	Sequence 1984, Ap
C 40	67.4	3.3	5748	16	US-10-311-455-1115	Sequence 1115, Ap
C 41	67.4	3.3	18154	16	US-10-311-455-228	Sequence 228, App
C 42	67.2	3.3	5807	16	US-10-311-455-1128	Sequence 1128, Ap
C 43	67.2	3.3	6816	21	US-10-723-860-7676	Sequence 7676, Ap
C 44	67	3.3	2053	14	US-10-027-632-97533	Sequence 97533, A
C 45	67	3.3	2053	18	US-10-027-632-97533	Sequence 97533, A

ALIGNMENTS

RESULT 1
US-10-473-126-386
; Sequence 386, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473.126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 386
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-386

Query Match 4.8%; Score 99.4; DB 21; Length 8056;
Best Local Similarity 44.5%; Pred. No. 1.5e-07;
Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;
Qy 548 ATGCTTCGAAATGTTTTTATATGTAATAATATGATCCCAAGGATAGTAATTC 607
Db 1975 ATTAATTAATTTTATTTTAAATTTTAAATATTTTAAATATTTTAAATATATAT 2034


```

Qy 986 GTCTCTGGCTGTCATTATATATAGGATCACTTACTGATCATCCATTAAAAACCTTGTTAA 1045
Db 1714246 ATATAAAATATATAAAATATATAAAATATATATAAAATATATAAAATATATAAA 1714187
Qy 1046 AACAAATTCATGAGATAAAATATCT-TACAAATGAAAGAGGACAATGTCTCTTTGAAA 1104
Db 1714186 ATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1714127
Qy 1105 AAACAAATAGTACTCTCCCTCGCTGAAATGTATACATATGGATTGGACACGGAGA 1164
Db 1714126 TATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1714067
Qy 1165 CTAGAAAAATGTATA-AAAGTAATGTAGAGTAAAGAGAGAGAGAGAGAGAGAGAG 1223
Db 1714066 ATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1714007
Qy 1224 GTACGGGACCCCAATATATAATTTGATAGATTTAGAAAGTAGTGGAAGTAGTGGGT 1283
Db 1714006 TATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1713947
Qy 1284 GGGTGGGATTTTATATATATAAAATTTTACTATTTTTCGAGAAAGTTTTCGAATGTATAGAA 1343
Db 1713946 ATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1713887
Qy 1344 TTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAATTAATGGGACAGAGGGAGTAA 1403
Db 1713886 ATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1713827
Qy 1404 CCTTATGATATATAAAATTTTGGTATTTTGGTATTTTCAATAGATTTAAATCTAGTTATA 1463
Db 1713826 TATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1713767
Qy 1464 ATGATAATATAATTTTAAATATACTATATTAAT---TCTGATAGTCGATTACCGC 1519
Db 1713766 TATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1713707
Qy 1520 CTTTATAAATTTTACAATCTGAGTAATATGAATATAATCACTTATCTGAAAGCAAAATA 1579
Db 1713706 AATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAA 1713647
Qy 1580 TATCTTTGTAAGACGGTTCGGTCAAAAGTGGAGTCACTGTGTATCAATAGTTTAAAT 1639
Db 1713646 AAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATA 1713587
Qy 1640 ATAAAGTAAATTTTAAATTAATTTTATTTTCTTTCAGAAATTTAAATTAATTTATG 1699
Db 1713586 ATAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATA 1713527
Qy 1700 AGCATGGGAAGTTTCACGGGATCATTCGACGACCTAGACTGTTTGAACAATGTATGTCC 1759
Db 1713526 AATTATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATA 1713467
Qy 1760 GGTGTACATCTAGACCTTCAACTCAACTCAAGTGAATATGCAATCTAGAAATACATCTT 1819
Db 1713466 ATAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATA 1713407
Qy 1820 TTCAAAATTCACAAACACAGCTTTTAACTTTTCTTTCACGGATTTGGAATCTTTCTAA 1879
Db 1713406 TAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATA 1713347
Qy 1880 ACTTTTAAAAATAAAAAA 1898
Db 1713346 AAAATATATAAAATATA 1713328

```

```

RESULT 3
US-10-473-126-240
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell

```

```

; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473.126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 240
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-240

```

```

Query Match 4.5%; Score 91.8; DB 21; Length 8056;
Best Local Similarity 44.2%; Pred. No. 3.4e-06;
Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;

Qy 547 AATGCTTCTCAAAATGTTTTTATATGTAATAAATGCCCCTCAAGGATTAAGTAAATTT 606
Db 1358 ATTTTATTAATAAATTTATTTTATATAAAATAATATCGATTTTTTTTTTATTTT 1417
Qy 607 CCCGTTTAAACAGCTTTGTTATATATATGTTTACCTTACAAGGATTAATTCGTAATCT 666
Db 1418 TTTATTTTTTTTTTAAAAAATAAATAAATTTATTTTTTAAATAAATATATATATAA 1477
Qy 667 TTTAGACGACAAAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCCTAGACTTGGTCA 726
Db 1478 ATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537
Qy 727 CTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTCAATAGCATTTAAATTTAG 786
Db 1538 TTAATTAATTAATTTTAAAAAATAATATTTTAAATTTAAATTTTAAATTTTAAATTTT 1597
Qy 787 GCTATTAATTAAGTTACTTAATAATAGAGAGGTTAGTAACAGAAACGAGTAAATTTT 846
Db 1598 TAAATTTAATTTTAAAAAATCGAAATAAACAAGTAAATTTAAAAAATAATTTT 1657
Qy 847 AGAGCTTGTGCTGTGTGTTAGTTGTGTGAGCTCATTTCTTTAAAAAGTAATGTAAACT 906
Db 1658 AATTAATAC-GATAAATAATTTTATTTTATTAATAATTAATAAATAATTAATTAATA 1716
Qy 907 GATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTCAAGAAATTTATATTTAA 966
Db 1717 TATATATTTTAAATTTTAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATA 1776
Qy 967 GAAATCATTTTATAACATGCTCTCGCTGTCTATTAATATAGGATCAGTCTACTGATCA 1026
Db 1777 AATTAATAAATAATTAATTAATTTTAAATTTTAAATAAATAAATAAATAAATAAATA 1836
Qy 1027 TCCATTAATAACCTTGTAAAAACAAATTCATAGATATAAATAATCTTACAAATGAAAA 1086
Db 1837 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1896
Qy 1087 GACAATGCTCTTTGAAAAACAAATAGGTACTCCCTCCGCTCCCTCTGGAATGTATACAT 1146
Db 1897 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1956
Qy 1147 ATGGAATGGACACGGACACTAAGAAAAATGTATAAAGTAAATGTAGAGTAAAAAGAGAG 1206
Db 1957 AATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2016
Qy 1207 A-----AAGAAAGTGGGTAAAGTAGTAGCGGACCCACCAATATATAATTTGATAGT 1256
Db 2017 ATATTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2076
Qy 1257 TTAGAAAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAATAAATTTACTAT 1316
Db 2077 ATTTTAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2136
Qy 1317 TTTGAGAAAGTTTGAATGTATAGATTTGATGGGACATCCATAAAGGAAAGTGTATA 1376
Db 2137 TTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2196

```


; SEQ ID NO 240
 ; LENGTH: 8056
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-473-126-240

Query Match 4.3%; Score 88.6; DB 21; Length 8056;
 Best Local Similarity 45.4%; Pred. No. 1.3e-05;
 Matches 442; Conservative 0; Mismatches 519; Indels 12; Gaps 3;

```

Qy 736 AGATAATGTTAGTATATATAGTAGGATCTACATCAGATATAAATTTAGAGCTATTAAT 795
Db 1399 ATATTATTTTATAAAAAAATAAATTTTAAATAAAAAATTAATAATTTAAATTTATTTT 1340
Qy 796 TAAGTTACTAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAGAGCTTGC 855
Db 1339 TTATTAAAAATTAATAAATAATTTTATATAATTTTAAATTTTAAAAAATAATTACGTTT 1280
Qy 856 TGCTGTGTTTGTGTTGTGAGCTCATTTCTTTAAAAAGTAATGTAAACCTGATCTAAAG 915
Db 1279 TTATATATTAATAATTAATCGTAACGTTTTTTTATTCGAAAAAATTTATTTATTA 1220
Qy 916 CACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTTATTTAAACGAAAAATCAT 975
Db 1219 AAAAAATTTAAATTAATAATTAATAATTTTAAAAATAAATTAATAATTCGAAAAATATA 1160
Qy 976 TTTATAACATGCTCTCGGCTGTCATTTATATAGGGATCAGTTACTGATCATCTCAATTA 1035
Db 1159 TTAATTTTAAATTTTCGAAATATCGAAAAAATAATTTAAAAACGAAATATATAACGTTTAA 1100
Qy 1036 ACCTTGTTTAAACCAAAATTCATAGATATAAATAATCTTTACAATGAAGAAGGCAATGTC 1095
Db 1099 AACGTTAAAAAATAAACAATTAACGAAATTTTATTTTACGTAATTAATTTTCGA 1040
Qy 1096 TCTTTGAAAAACAAATAGGTACTCCCTCGCTCCCTGAAATGTATACATATATGATGTTG 1155
Db 1039 AATTAAAAAATAAATAAATAATTTTAAAAATAATTTTAAAAAATAATTAATAATA 980
Qy 1156 ACAGGAGACTAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAAAGAAAG 1215
Db 979 TTTTATTCGTTTAAATAAATAAATAAATAATTTTATTTTAAAAAATAATTTTATTCGAAAC 920
Qy 1216 TGGTAAAGTAGCGGACCCACCAATATATAATTTGATAGATTTAGAAAAAGTGTGGAAG 1275
Db 919 GACGAATAATTAATAAATAAATAATTTAAATAATTTACGAAATTTTATACGTAATAA 860
Qy 1276 TAGTGGTGGGTGGGATTTTATATATAAAAAATTTACTAT-TTTGAGAAAGTTTGGAA 1334
Db 859 TAAAAAAACGATTTTATTAATAAATAATTTTAAATAATCGTTTATAAATTTATTTTA 800
Qy 1335 TGTATAGAAATGAGTGGGACATCCATAAAGGAAGGTGTAGAAATTAATTTGGACAGAG 1394
Db 799 AATATATAATTAATAATAATTTTATATAATTTTAAATTTTAAAAATTTAAAAATTTAAT 740
Qy 1395 GGAGTAATACCTTTATCATATATAAATTTTGTGTAT-----TTTGATTTTCAATAAGTTA 1448
Db 739 TATAATAAATAATTTTAAATTTAAATTTTAAATAAATAAATAATTTTATATTTATTTA 680
Qy 1449 TAAATCTATGTATTAATAATAATAATTTTAAAAATAAATACTATATTAATTTCTGATTAG 1508
Db 679 TTTTATTTTATTTTAAATAAATTTATATTTTAAAAATTTTAAAAAATTTTAAAAAT 620
Qy 1509 TCGATTACCGCTTTTATAA-----TTTTCATACTGAGTAATATGAATAAATTCAGTTA 1563
Db 619 TATATAATAATTTTATATAAATAATTTATTAATAATAAATAAATTTTAAATAAACAAGTTTA 560
Qy 1564 TCTGAAAGCAATAATATCTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAGTTTCATGTT 1623
Db 559 TAATAATATATAAATAATTTTAAATTTTAAATATATAAATAATTTTAAATATTAATAATA 500
Qy 1624 ATTCAATAGTTTTTAATAAATAAAGTAAATTTTAAATTAATTTGTTATTTTCTGTTT 1683
  
```

RESULT 6

US-10-211-179-11

; GENERAL INFORMATION:

; APPLICANT: Nicholas M. Dean

; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTYROSYL PHOSPHATASE ACTIVATOR EXP

; FILE REFERENCE: PFS-0011

; CURRENT APPLICATION NUMBER: US/10/211,179

; CURRENT FILING DATE: 2002-08-01

; NUMBER OF SEQ ID NOS: 119

; OTHER INFORMATION: n = a, t, c, or g

US-10-211-179-11

Query Match 4.0%; Score 83; DB 18; Length 158001;

Best Local Similarity 46.0%; Pred. No. 0.00035;

Matches 458; Conservative 0; Mismatches 530; Indels 7; Gaps 5;

```

Qy 708 AACAGCCTAGACTTTGGTCACTGATAAATAGATATATTCCTTAGTATAATATAGTAGGATCTA 767
Db 126953 AATATATTATATATATATATATATATATATATATATATATATATATATATATATATATA 127012
Qy 768 CAATGACATTTAAATTTAGAGCTATTAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAA 827
Db 127013 ATAATATATAATATAATAATATATATATATATATATATATATATATATATATATATAA 127072
Qy 828 CAGAAAGCAGGTAACAAACAGAGCTTCCTGCTGCTGTTAGTTGTTGTCGAGCTCATTTTC 887
Db 127073 AAAATAATATATAAATAATATAAATAATATATAAATAATATATAAATAATATATAAATA 127132
Qy 888 TTTAAAGAGTAATCTAAACCTGATCTAAAGCACATAGA-AAATTAGTACAGGTTAAAACTTT 946
Db 127133 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 127192
Qy 947 TACAAGAATTTATATTAACGAAAAATCATTTTATAACATGCTCTCGGCTGTCAATTATAA 1006
Db 127193 TAAAAATATATATATATATATATATATAAATAATATAAATAATATAAATAATATATATA 127252
Qy 1007 TAGGGATCACTTACTGATCATTCATTAACCTTTGTTAAAAACAAATTCATAGAGATAAAA 1066
Db 127253 AATATATAAATAATATAAATAATATATATATAAATAATATAAATAATATAAATAATATA 127312
Qy 1067 TATCTTTACAATG-AAAAGAGGACAATGCTCTTTGAAAAAACAAATAGGTACTCCCTCC 1125
Db 127313 TATATAAATATATAAATAATATATAATATATATATATATATAAATAATATAAATAATAT 127372
Qy 1126 GTCCCTCTGAAATGTATACATATGGATTGGACGCGAGACTAAAGAAAAATGTATPAAAGTA 1185
Db 127373 ATATTATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATATAAATA 127432
Qy 1186 ATGTAGAGTAAAAAGAAAGAAAGTGGGTAAAGTGGGAGCCGCCACCAATATAT 1245
Db 127433 TAATATATATATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATAT 127492
Qy 1246 AATTGTATAGATTTTAGAAAAAGTCTGAAAGTCTGGGTGGGTGGGATTTTATATATATAA 1305
Db 127493 AATATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAATAT 127552
Qy 1306 AAATTTACTATTTTGGAAAAAGTTTTCGAAATGTATAGAAATTTGAGTGGGACATCCATAAAG 1365
Db 127553 AAAAAATATACAATATATAAATAATATAAATAATATAAATAATATAAATAATATAAATA 127612
Qy 1366 GAAAGTGTATAGAAATTTAAATGGGACAGAGGGAGTA-ATACCTTTATGATATATAAATTTT 1424
Db 127613 TATATAATATATAAATAAATAATACAAATATATATAAATAATATAAATAATATAAATAATAT 127672
  
```


; CURRENT APPLICATION NUMBER: US/10/240,453
 ; CURRENT FILING DATE: 2002-10-02
 ; PRIORITY APPLICATION NUMBER: PCT/EP01/03973
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 350
 ; SEQ ID NO 206
 ; LENGTH: 11745
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (9105)
 ; US-10-240-453-206

 Query Match 3.7%; Score 75; DB 16; Length 11745;
 Best Local Similarity 42.7%; Pred. No. 0.0038;
 Matches 499; Conservative 0; Mismatches 665; Indels 5; Gaps 2;

 QY 768 CAATGACATTAAATTAGAGCTATTAAATTAAAGTTACTTAATAATAAGAGAGGTTAGTAAA 827
 DB 4805 CGAATAAATAAATAATCTTATTTCTAAATAAAAAAATTAATAAAAAAACAATTAATTT 4746
 QY 828 CAGAAACAGGTAAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
 DB 4745 AATAAATAAATAAATAATCTTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4686
 QY 888 TTTAAAGTAAATGAAAGTAAATGATCTAAAGCAGCATAGAAATTTAGTACAGGTTAAACCTTTT 947
 DB 4685 CTCATTACACATCTACCTAAATAATTTAAATAAACCACTATATATATAATACTTAAACCTTA 4626
 QY 948 ACAAGAATTTATATAACGAATCATTTTATTAACATGCTCTCGGCTGCTCATTAAT 1007
 DB 4625 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4566
 QY 1008 AGGATCAGTCTTACTGATCATCTTAAACCTTGTAAACCAATTCGAATGAGTAAAT 1067
 DB 4565 ACCTTAAATTAATCTGCAATCACCCTAAAA--AAAAAACAATAAATAAATAAATAAATAA 4509
 QY 1068 ATCTTACAATGAAAGAGGACATGCTCTCTTTGAAACCAATAAGTACTCCCTCCGT 1127
 DB 4508 ACCAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4449
 QY 1128 CCCTCTGAAATGATACATATGGAATGGAACGAGACTAAGAAAAATGTTAAGTAAT 1187
 DB 4448 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4389
 QY 1188 GTAGAGTAAAAAGAGAGAGAAAGTGGTAAAGTAGCGGACCCCAATATATAA 1247
 DB 4388 ATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4331
 QY 1248 TTGATAGATTTAGAAAAAGTAGTTGAAAGTAGTGGGTTGGGATTTTATATATATAA 1307
 DB 4330 TTTAAATCCTAAACCAAACTAAATAATCATTTTAAATTTTATCAATATATACTCAATAAT 4271
 QY 1308 ATTTACTATTTTGAGAAAGTTTGAATGTTATGAAATGTTAGTGGGACATCCATAAAGGA 1367
 DB 4270 TATTTTAAATTAATTTAAACCTCTTTTATTAATAAATAAATAAATAAATAAATAAATAA 4211
 QY 1368 AAGTGATAGAAATTAATGGACAGAGGAGTATACCTTTATGATATATAAATTTTGT 1427
 DB 4210 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4151
 QY 1428 TATTTGATTTCAATGATTAATAATCTATGTTTATAATGATATAATTTTAAATAA 1487

DB 4150 TTCTACTAAATAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4091
 QY 1488 TACTATATTAATTTCTGATTTAGTTCGATTTACGCGCTTTTATATATTTTACAACTAGTGAAT 1547
 DB 4090 TTTTCTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4031
 QY 1548 ATGAATAAATCAGTTATCTCAAAAGCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1607
 DB 4030 AAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3971
 QY 1608 TGGGAAGTTCATGTTTCAATAGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1667
 DB 3970 AAACAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3911
 QY 1668 TTTTGTTCAGAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1727
 DB 3910 AAAAATTTTCAATTTTAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3851
 QY 1728 GCAGCACTAGACTGTTTGAACAATGATGTCGCGTGACATCTATGACCTTTTCAACTCAA 1787
 DB 3850 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3791
 QY 1788 ACTAGTGAATAATGCAATTTCTAGATAATACATCTTTTCAAAATTTCAACAAACAGCTTTAAC 1847
 DB 3790 AACATAAAACAATAATTCATCAACCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3731
 QY 1848 TTTTCTTTCAACGATTTGGAATCCTTTTCTAAACTTTTAAATAAATAAATAAATAAATAAATAA 1907
 DB 3730 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3671
 QY 1908 TGTAAATATTTATCAACACCTCAACATTTGA 1936
 DB 3670 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3642

RESULT 9

US-10-981-277-34/c
 ; Sequence 34, Application US/10981277
 ; Publication No. US20050181389A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harris, Cole
 ; TITLE OF INVENTION: Compositions and Methods for Glioma Classification
 ; FILE REFERENCE: 03-968-US
 ; CURRENT APPLICATION NUMBER: US/10/981,277
 ; PRIOR FILING DATE: 2004-11-03
 ; PRIOR APPLICATION NUMBER: US 60/516,817
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 34
 ; LENGTH: 176096
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-981-277-34

Query Match 3.5%; Score 71.8; DB 24; Length 176096;
 Best Local Similarity 48.9%; Pred. No. 0.036;
 Matches 280; Conservative 0; Mismatches 287; Indels 6; Gaps 3;

 QY 1135 AAATGTATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAGTAATCTAGAGT 1194
 DB 167242 AAATTTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 167183
 QY 1195 AAAAAAGAGAGAGAAAAAGTCGGTAAAGTAGCGGACCCCAACAATATATAAATTGATAG 1254
 DB 167182 ATAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 167125
 QY 1255 ATTTAGAAAGTAGTTGAAAGTAGTGGTGGGATTTTATATATATAAATAAATAAATAA 1314
 DB 167124 ATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 167065

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2005, 18:42:01 ; Search time 9010 Seconds
(without alignments)
11035.528 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcaatacattctaaatc 2052

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues
Word size : 50

Total number of hits satisfying chosen parameters: 15

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2052	100.0	2052	6 E40087	E40087 plant promo
2	2037	99.3	2052	6 E40093	E40093 plant promo
3	1809	88.2	2056	6 E40091	E40091 plant promo
4	1748	85.2	2048	6 E40090	E40090 plant promo
5	1733	84.5	2048	6 E40089	E40089 plant promo
6	57	2.8	140	6 AR146856	AR146856 Sequence
7	57	2.8	140	6 AR146857	AR146857 Sequence
8	57	2.8	140	6 E55069	E55069 plant promo
9	57	2.8	140	6 E55070	E55070 plant promo
10	57	2.8	246	6 AR146852	AR146852 Sequence
11	57	2.8	246	6 E55065	E55065 plant promo
12	57	2.8	247	6 AR076816	AR076816 Sequence
13	57	2.8	247	6 E15124	E15124 Promoter. 7
14	57	2.8	2042	6 AR076817	AR076817 Sequence
15	57	2.8	2042	6 E15125	E15125 Promoter. 7

ALIGNMENTS

```
RESULT 1
E40087
LOCUS      E40087          2052 bp          DNA          linear          PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION  E40087
VERSION    E40087.1  GI:18627203
KEYWORDS   JP 2000166577-A/1.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 2052)
AUTHORS   Nishikawa,S. and Oeda,K.
TITLE      Plant promoter and terminator
JOURNAL    Patent: JP 2000166577-A 1 20-JUN-2000;
           SUMITOMO CHEM CO LTD
COMMENT    OS Daucus carota L.
           PN JP 2000166577-A/1
           PD 20-JUN-2000
           PF 01-OCT-1999 JP 1999281475
           PR
           PI SATOMI NISHIKAWA,KENJI OEDA
           PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//(C12N5/10,C12R1:91), PC
           C12N15/00,
           PC C12N5/00,(C12N5/00,C12R1:91)
           CC
           FH Key Location/Qualifiers
           FT promoter (1)..(2052).
           Location/Qualifiers
           1..2052
           /organism="unidentified"
           /mol_type="genomic DNA"
           /db_xref="taxon:32644"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 2052; DB 6; Length 2052;
Best-Local Similarity 100.0%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CATGTGCGCTACAGCACATAGGCGCTGTTGGTTGAGAGAGACAGAGCTGTTCTGA 60
Db 1 CATGTGCGCTACAGCACATAGGCGCTGTTGGTTGAGAGAGACAGAGCTGTTCTGA 60
Qy 61 CTTCTCTCTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTTAAAGCTGCGAA 120
Db 61 CTTCTCTCTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTTAAAGCTGCGAA 120
Qy 121 TACTAACTTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTTAACTTTTACT 180
Db 121 TACTAACTTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTTAACTTTTACT 180
Qy 181 TCTCATTTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTTAACTTTTACT 240
Db 181 TCTCATTTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTTAACTTTTACT 240
Qy 241 AACGGCTCAATAAAGATCATTTCAATAATGTAATCTTTCAATTTTAGGATAACAATACGT 300
Db 241 AACGGCTCAATAAAGATCATTTCAATAATGTAATCTTTCAATTTTAGGATAACAATACGT 300
Qy 301 GAA CAGGGTTATTTTAACTGTCACAAATTTCTAATAATTTTAACTGCGCGGTGAACA 360
Db 301 GAA CAGGGTTATTTTAACTGTCACAAATTTCTAATAATTTTAACTGCGCGGTGAACA 360
Qy 361 CCGTCTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAACTGCGGTGTC 420
Db 361 CCGTCTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAACTGCGGTGTC 420
Qy 421 AGACAGCTTGTAGTGAATACACATTTGACGTCTTTTAAACAAAGAACAGTGGTTC 480
Db 421 AGACAGCTTGTAGTGAATACACATTTGACGTCTTTTAAACAAAGAACAGTGGTTC 480
Qy 481 ATGCTCAGCATCAAAATTCAGAAAACCCGACACACACTCTATCCAGTACTATCTTT 540
Db 481 ATGCTCAGCATCAAAATTCAGAAAACCCGACACACACTCTATCCAGTACTATCTTT 540
```

481	Db	ATGCTCAGCCATCAAAATTGACAAAAACCGGACACAACTCTATCCAGCTACTATACTTT	540
541	Qy	TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAAAATAGTCCCATCCAAGGATAAGT	600
541	Db		
541	Db	TGGCCGAATGCTTCTCAAAATGTTTTTTATATGTAAAAAATAGTCCCATCCAAGGATAAGT	600
601	Qy	AAAAATCCGGTTTAAACAGTTTGTTAAATATATATGTTTTACACTTACAAGAGGATATTCGT	660
601	Db		
601	Db	AAAAATCCGGTTTAAACAGTTTGTTAAATATATATGTTTTACACTTACAAGAGGATATTCGT	660
661	Qy	AATACTTTTATAGACACAGAGACTTAGTGCAAAATCGACGCTGGTAAACAGCCTAGACT	720
661	Db		
661	Db	AATACTTTTATAGACACAGAGACTTAGTGCAAAATCGACGCTGGTAAACAGCCTAGACT	720
721	Qy	TGGTCACTGATAAATAGATAAATCTTTAGTATAATATAGTAGGATCTACAATGACATTTAAA	780
721	Db		
721	Db	TGGTCACTGATAAATAGATAAATCTTTAGTATAATATAGTAGGATCTACAATGACATTTAAA	780
781	Qy	ATTAGAGCTATTAAATTAAGTTACTAAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTA	840
781	Db		
781	Db	ATTAGAGCTATTAAATTAAGTTACTAAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTA	840
841	Qy	AAAAACAGAGCTGCTGCTGTGTGTTAGTGTGTGAGCTCATTTCTTTAAAAAGTAATG	900
841	Db		
841	Db	AAAAACAGAGCTGCTGCTGTGTGTTAGTGTGTGAGCTCATTTCTTTAAAAAGTAATG	900
901	Qy	TAAACTGATCTAAAGACACATAGAAAATTTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA	960
901	Db		
901	Db	TAAACTGATCTAAAGACACATAGAAAATTTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA	960
961	Qy	TTAAACGAAAATCAATTTTATAACATGCTCTCTCGGCTGTCATTATAATAGGGATCAGTTAC	1020
961	Db		
961	Db	TTAAACGAAAATCAATTTTATAACATGCTCTCTCGGCTGTCATTATAATAGGGATCAGTTAC	1020
1021	Qy	TGATCATCCATTAAACCTTGTTTAAAAACAAATTCAAATGAGATATAAATATCTTACAATGAA	1080
1021	Db		
1021	Db	TGATCATCCATTAAACCTTGTTTAAAAACAAATTCAAATGAGATATAAATATCTTACAATGAA	1080
1081	Qy	AAGAAGGACAATGTCCTTTTGAAAAACAAATAGGTACTCCCTCCGCTCCCTCTGAAAATGT	1140
1081	Db		
1081	Db	AAGAAGGACAATGTCCTTTTGAAAAACAAATAGGTACTCCCTCCGCTCCCTCTGAAAATGT	1140
1141	Qy	ATACATATGGATTGGACACGCGAGACTTAAGAAAAATGTATAAGTAAATGTAGAGTAAAAAG	1200
1141	Db		
1141	Db	ATACATATGGATTGGACACGCGAGACTTAAGAAAAATGTATAAGTAAATGTAGAGTAAAAAG	1200
1201	Qy	AAAGAGAAAGAAAAGTGCGTAAAGTAGCGGACCCACAAATATAATAATGTATAGATTAG	1260
1201	Db		
1201	Db	AAAGAGAAAGAAAAGTGCGTAAAGTAGCGGACCCACAAATATAATAATGTATAGATTAG	1260
1261	Qy	AAAAGTAGTCAAAAGTAGTGGGTGGGATTTTTTATATATATAAAAATTTTACTATTTTG	1320
1261	Db		
1261	Db	AAAAGTAGTCAAAAGTAGTGGGTGGGATTTTTTATATATATAAAAATTTTACTATTTTG	1320
1321	Qy	AGAAAGTTTTGAAATGTATAGAAATGTAGTGGGACATCCATAAAAAGGAAAGTGTATAGAAT	1380
1321	Db		
1321	Db	AGAAAGTTTTGAAATGTATAGAAATGTAGTGGGACATCCATAAAAAGGAAAGTGTATAGAAT	1380
1381	Qy	TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAAATTTTGTGTTATTTTGATTTCA	1440
1381	Db		
1381	Db	TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAAATTTTGTGTTATTTTGATTTCA	1440
1441	Qy	TAGATTATAAATCTATGTTTATAATGATAATATAATTTTAAAAATACTATATTTAAAT	1500
1441	Db		
1441	Db	TAGATTATAAATCTATGTTTATAATGATAATATAATTTTAAAAATACTATATTTAAAT	1500
1501	Qy	CTGATTAGTCGATTACCGCCCTTTTATAATTTTACAATCTAGTAAATATGAATAAATCAG	1560
1501	Db		
1501	Db	CTGATTAGTCGATTACCGCCCTTTTATAATTTTACAATCTAGTAAATATGAATAAATCAG	1560
1561	Qy	TTATCTGAAAAAGCAAAATAATCTTTGTGTAACAGCGCTTCGGTCAAAATGGGAAGTTCATG	1620
1561	Db		
1561	Db	TTATCTGAAAAAGCAAAATAATCTTTGTGTAACAGCGCTTCGGTCAAAATGGGAAGTTCATG	1620

1621	Qy	TGTAATCAATAGATTTAAATAAAAGTAAATTTAAATTAATTTGTTTATTTTGGTTTCAGA	1680
1621	Db	TGTAATCAATAGATTTAAATAAAAGTAAATTTAAATTAATTTGTTTATTTTGGTTTCAGA	1680
1681	Qy	AAATTTAAAAATAATTAATGAGCATGGGAAGTTACACGGGCATCATTGAGCAGCACTAGACT	1740
1681	Db	AAATTTAAAAATAATTAATGAGCATGGGAAGTTACACGGGCATCATTGAGCAGCACTAGACT	1740
1741	Qy	GTTTGAACAATGTATGTCCGGGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT	18000
1741	Db	GTTTGAACAATGTATGTCCGGGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT	18000
1801	Qy	GCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACACAGCTTTTAACTTTTCTTTCAACG	1860
1801	Db	GCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACACAGCTTTTAACTTTTCTTTCAACG	1860
1861	Qy	GATTGGAATCCTTTTCTAAACTTTTAAAAATAAAAAAAATGCAATTAATGTAATATTTATC	1920
1861	Db	GATTGGAATCCTTTTCTAAACTTTTAAAAATAAAAAAAATGCAATTAATGTAATATTTATC	1920
1921	Qy	AACACCTCAACATGTATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT	1980
1921	Db	AACACCTCAACATGTATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT	1980
1981	Qy	CACATCAATCTTACACCAACAAACCTTGAGCTTAATTTTCTACTTAATTTCTCAGCAATAAAC	2040
1981	Db	CACATCAATCTTACACCAACAAACCTTGAGCTTAATTTTCTACTTAATTTCTCAGCAATAAAC	2040
2041	Qy	ATTCTAAATATC 2052	
2041	Db	ATTCTAAATATC 2052	

RESULT 2	E40093	2052 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	E40093				
DEFINITION	Plant promoter and terminator.				
ACCESSION	E40093				
VERSION	E40093.1 GI:18627209				
KEYWORDS	JP 2000166577-A/7.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 2052)				
AUTHORS	Nishikawa, S. and Oeda, K.				
TITLE	Plant promoter and terminator				
JOURNAL	Patent: JP 2000166577-A 7 20-JUN-2000;				
COMMENT	SUMITOMO CHEM CO LTD				
	OS Daucus carota L.				
	PN JP 2000166577-A/7				
	PD 20-JUN-2000				
	PF 01-OCT-1999 JP 1999281475				
	PR				
	PI SATOMI NISHIKAWA, KENJI OEDA				
	PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC C12N15/00,				
	PC C12N5/00, (C12N5/00, C12R1:91)				
	CC				
FEATURES	Key	Location/Qualifiers			
source	promoter	(1)..(2052).			
	1..2052	Location/Qualifiers			
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	99.3%;	Score 2037;	DB 6;	Length 2052;	
Best Local Similarity	100.0%;	Pred. NO. 0;			
Matches 2037;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QV	1	CATGTGTGCCCTACAGCACATAGGCGCTGTGTTGGTTGAGAGAGACAGCAAGCTGCTTCTGA	60		

Db 1 CATGTGCGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGCAAGCTGCTCTGA 60
Qy 61 CTTCTCTCTTTTGACCTGTTGTATAAAGAGTGAATAATTTTTAAAAAGCTGCGAA 120
Db 62 CTTCTCTCTTTTGACCTGTTGTATAAAGAGTGAATAATTTTTAAAAAGCTGCGAA 120
Qy 121 TACTAACTCTCTCTCACAACTCCGCTCTCTTTTCCAAACACATTTATTAACCTTTTACT 180
Db 121 TACTAACTCTCTCTCACAACTCCGCTCTCTTTTCCAAACACATTTATTAACCTTTTACT 180
Qy 181 TCTCATTTCTCTCCACTCTTTGCTATAAGCAAGAAATCACTCTCTTTTAAAGCTAACCCA 240
Db 181 TCTCATTTCTCTCCACTCTTTGCTATAAGCAAGAAATCACTCTCTTTTAAAGCTAACCCA 240
Qy 241 AACGGCCCTCAATAAAGATCAATCAATAAATGTAATCTTTCAATTTTAGGATAACAATACGT 300
Db 241 AACGGCCCTCAATAAAGATCAATCAATAAATGTAATCTTTCAATTTTAGGATAACAATACGT 300
Qy 301 GAACAGGGTTATTTTTAAAGTGTCAACAAATCTTAATAATTTTACCTGGCGGTGAACA 360
Db 301 GAACAGGGTTATTTTTAAAGTGTCAACAAATCTTAATAATTTTACCTGGCGGTGAACA 360
Qy 361 CCGTCTTCCAAGATAATATTTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Db 361 CCGTCTTCCAAGATAATATTTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Qy 421 AGGACGACTTAGGTGAATACACATTTGTAAGTGTCTTTTAAACCAAGAGCAAGTGGTTC 480
Db 421 AGGACGACTTAGGTGAATACACATTTGTAAGTGTCTTTTAAACCAAGAGCAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGAACAAACCCGACACAACTCTATCCACGTAATCTACTTT 540
Db 481 ATGCTCAGCCATCAAAATTTGAACAAACCCGACACAACTCTATCCACGTAATCTACTTT 540
Qy 541 TGGCCGAATGCTCTCAAAATGTTTTTATATATATATATATATATATATATATATATATAT 600
Db 541 TGGCCGAATGCTCTCAAAATGTTTTTATATATATATATATATATATATATATATATATAT 600
Qy 601 AATAATCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATATAT 660
Db 601 AATAATCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATATAT 660
Qy 661 AATACTTTTACGACACAGAGACTTAGGTCAAAATCGACGCTGTTAAACAGCCTAGACT 720
Db 661 AATACTTTTACGACACAGAGACTTAGGTCAAAATCGACGCTGTTAAACAGCCTAGACT 720
Qy 721 TGCTCACTGATAAATAGATAATTTGTATATATATATATATATATATATATATATATATATAT 780
Db 721 TGCTCACTGATAAATAGATAATTTGTATATATATATATATATATATATATATATATATATAT 780
Qy 781 ATTAGAGCTATTAATTAAGTTACTTAATAATAAGAGAGTTAGTAAACAGAAAGCAGGTA 840
Db 781 ATTAGAGCTATTAATTAAGTTACTTAATAATAAGAGAGTTAGTAAACAGAAAGCAGGTA 840
Qy 841 AAAACAAGAGCTCTGCTGCTGTTAGTTGTTGTCAGCTCATTTCTTTAAAAAGTAATG 900
Db 841 AAAACAAGAGCTCTGCTGCTGTTAGTTGTTGTCAGCTCATTTCTTTAAAAAGTAATG 900
Qy 901 TAAACTGATCTAAAGCACAAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA 960
Db 901 TAAACTGATCTAAAGCACAAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA 960
Qy 961 TTAACGAAATCAATTTTAAACATGCTCTCGGCTGCTCATTATTAATAGGATCAGTTAC 1020
Db 961 TTAACGAAATCAATTTTAAACATGCTCTCGGCTGCTCATTATTAATAGGATCAGTTAC 1020
Qy 1021 TGATCATCTCAATTAACCTCTGTTAAACAAATTCATAGAGATAAATATCTTACAATGAA 1080
Db 1021 TGATCATCTCAATTAACCTCTGTTAAACAAATTCATAGAGATAAATATCTTACAATGAA 1080
Qy 1081 AAGAGGACATGCTCTCTTTGAAAAACAAATAGGTACTCCCTCCGCTCTCGAAATGT 1140

Db 1081 AAGAGGACAAATGCTCTCTTTGAAAAACAAATAGGTACTCCCTCCGCTCTCGAAATGT 1140
Qy 1141 ATACATATGATTTGGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1200
Db 1141 ATACATATGATTTGGACACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1200
Qy 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAATATATATAATTTGATAGATTTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAATATATATAATTTGATAGATTTAG 1260
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1320
Qy 1321 AGAAAGTTTTGAAATGATATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
Db 1321 AGAAAGTTTTGAAATGATATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
Qy 1381 TAAATGGGACAGAGGGAGTAATACCTTTTATGATATATAAATTTTGTATTGTTGATTCTCA 1440
Db 1381 TAAATGGGACAGAGGGAGTAATACCTTTTATGATATATAAATTTTGTATTGTTGATTCTCA 1440
Qy 1441 TAAGATTAATAATCTATGTTTATTAATGATAATATAATTTTAAAAATAATATATATTAAT 1500
Db 1441 TAAGATTAATAATCTATGTTTATTAATGATAATATAATTTTAAAAATAATATATATTAAT 1500
Qy 1501 CTGATTAGTCGATTACCGCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1560
Db 1501 CTGATTAGTCGATTACCGCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1560
Qy 1561 TTATCTGAAAGCAAAATAATCTTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAAGTTCA 1620
Db 1561 TTATCTGAAAGCAAAATAATCTTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAAGTTCA 1620
Qy 1621 TGATTTCAATAGTTTTTAATAATAAAGTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1680
Db 1621 TGATTTCAATAGTTTTTAATAATAAAGTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1680
Qy 1681 AATTTAAAAATAAATTTATGAGCATGGGAGTTACCGGGCATCTTGTAGAGCAGCAGTACT 1740
Db 1681 AATTTAAAAATAAATTTATGAGCATGGGAGTTACCGGGCATCTTGTAGAGCAGCAGTACT 1740
Qy 1741 GTTTGAACAATGATGTCGCGGTGTACATCTATGACCTTTTCAAACTCAAACTAGTGAATAAT 1800
Db 1741 GTTTGAACAATGATGTCGCGGTGTACATCTATGACCTTTTCAAACTCAAACTAGTGAATAAT 1800
Qy 1801 GCATTTAGAAATACATCTTTTCAAAATTTTCAAAACACAGCGTTTAACTTTTCTTTCAACG 1860
Db 1801 GCATTTAGAAATACATCTTTTCAAAATTTTCAAAACACAGCGTTTAACTTTTCTTTCAACG 1860
Qy 1861 GATTGGAATCCTTTTCTTAAACTTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAA 1920
Db 1861 GATTGGAATCCTTTTCTTAAACTTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAA 1920
Qy 1921 AACACCTCAACATTTGATGTTAGCGGTACTATAAATAGGTGCTCTTGGTCTCTACTCAT 1980
Db 1921 AACACCTCAACATTTGATGTTAGCGGTACTATAAATAGGTGCTCTTGGTCTCTACTCAT 1980
Qy 1981 CACATCAATCTTACACCAAAACCTTGAGCTTAAATTTTCTACTTATTTCTCAGCAAT 2037
Db 1981 CACATCAATCTTACACCAAAACCTTGAGCTTAAATTTTCTACTTATTTCTCAGCAAT 2037

RESULT 3
E40091
LOCUS
DEFINITION Plant promoter and terminator.
ACCESSION E40091
VERSION E40091.1 GI:18627207
KEYWORDS JP 2000166577-A/5.
SOURCE unidentified
ORGANISM unclassified.

E40091 2056 bp DNA linear PAT 31-JAN-2002

REFERENCE	1 (bases 1 to 2056)	
AUTHORS	Nishikawa, S. and Oeda, K.	
TITLE	Plant promoter and terminator	
JOURNAL	Patent: JP 2000166577-A 5 20-JUN-2000;	
COMMENT	SUMITOMO CHEM CO LTD	
	OS Daucus carota L.	
	PN JP 2000166577-A/5	
	PD 20-JUN-2000	
	PF 01-OCT-1999 JP 1999281475	
	PR SATOMI NISHIKAWA, KENJI OEDA	
	PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC	
	C12N15/00,	
	PC C12N5/00, (C12N5/00, C12R1:91)	
CC	Key	Location/Qualifiers
FT	promoter	(1)..(2056).
FEATURES	Location/Qualifiers	
source	1..2056	
	/organism="unidentified"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32644"	
ORIGIN		
	Query Match	88.2%; Score 1809; DB 6; Length 2056;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 1809; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CATGTGTGCCCTACAGCACATAGGGCTGTGGTGGAGAGAGCAGAAGCTGCTCTGA 60
Db	1	CATGTGTGCCCTACAGCACATAGGGCTGTGGTGGAGAGAGCAGAAGCTGCTCTGA 60
Qy	61	CTTCTCTCTTTTGACCTGTGTATAAAGAGTAGAATAATTTTTAAAAAGTCGAA 120
Db	61	CTTCTCTCTTTTGACCTGTGTATAAAGAGTAGAATAATTTTTAAAAAGTCGAA 120
Qy	121	TACTAACTTCTCTCACAACCTCCGCTCTTTTCCAAAACACTTTATTAACCTTTTACT 180
Db	121	TACTAACTTCTCTCACAACCTCCGCTCTTTTCCAAAACACTTTATTAACCTTTTACT 180
Qy	181	TCTCATTTCTACTCCACTCTTTGCTATAGCAAGAAATCACTCTTTTAAAGCTTAACCCA 240
Db	181	TCTCATTTCTACTCCACTCTTTGCTATAGCAAGAAATCACTCTTTTAAAGCTTAACCCA 240
Qy	241	AACGGCTCAATAAAGATCAATCATAAATGTATCTTTCAATTTAGGATAACAATACGT 300
Db	241	AACGGCTCAATAAAGATCAATCATAAATGTATCTTTCAATTTAGGATAACAATACGT 300
Qy	301	GAACAGGGTTATTTTAAAGTGTCAACAAATCTTAATAATTTTACCTGGCCGGTGAACA 360
Db	301	GAACAGGGTTATTTTAAAGTGTCAACAAATCTTAATAATTTTACCTGGCCGGTGAACA 360
Qy	361	CCGCTCTTCCAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGCATGC 420
Db	361	CCGCTCTTCCAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGCATGC 420
Qy	421	AGGACGACTTAGGTGAATACACATGTGTAGTCTTTAAACAAAGAACAGTGGTTC 480
Db	421	AGGACGACTTAGGTGAATACACATGTGTAGTCTTTAAACAAAGAACAGTGGTTC 480
Qy	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCACGTACTATCTTT 540
Db	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCACGTACTATCTTT 540
Qy	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATGATAAATAATGCCCATCCAAAGATAAGT 600
Db	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATGATAAATAATGCCCATCCAAAGATAAGT 600
Qy	601	AAAAATCCGGTTAACAGTTTGTAAATATATATGTTTACACTTACAAGAGGATATTCGT 660
Db	601	AAAAATCCGGTTAACAGTTTGTAAATATATATGTTTACACTTACAAGAGGATATTCGT 660
Qy	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720

Db	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
Qy	721	TGCTCACTGATAAATAGATAAATTTGTAGTATAATATAGTAGGATCTACAATGACATAAA 780
Db	721	TGCTCACTGATAAATAGATAAATTTGTAGTATAATATAGTAGGATCTACAATGACATAAA 780
Qy	781	ATTAGAGCTATTAAATTAAGTTACTATAATAAGAGAGGTTAGTAAACAGAAACAGGTA 840
Db	781	ATTAGAGCTATTAAATTAAGTTACTATAATAAGAGAGGTTAGTAAACAGAAACAGGTA 840
Qy	841	AAAAAAGAGCTTGTGCTGTGTTTGTAGTTGTGTGAGCTCATTTCTTTAAAAAGTAAATG 900
Db	841	AAAAAAGAGCTTGTGCTGTGTTTGTAGTTGTGTGAGCTCATTTCTTTAAAAAGTAAATG 900
Qy	901	TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA 960
Db	901	TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA 960
Qy	961	TTAAAACGAAATCATTTTATAACATGCTCTCGGCTGTCAATTAATAGGATCAGCTTAC 1020
Db	961	TTAAAACGAAATCATTTTATAACATGCTCTCGGCTGTCAATTAATAGGATCAGCTTAC 1020
Qy	1021	TGATCATCTCAATTAATAACCTTTTAAACAAATTTCAATGAGATAAAATATCTTACAATGAA 1080
Db	1021	TGATCATCTCAATTAATAACCTTTTAAACAAATTTCAATGAGATAAAATATCTTACAATGAA 1080
Qy	1081	AAGAAGGACAAATGCTCTTTTGAATAAAACAAATAGTACTCCCTCCGCTCTGAAATGT 1140
Db	1081	AAGAAGGACAAATGCTCTTTTGAATAAAACAAATAGTACTCCCTCCGCTCTGAAATGT 1140
Qy	1141	ATACATATGATTTGGACACCGGACCTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Db	1141	ATACATATGATTTGGACACCGGACCTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Qy	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATATAATTTGATAGATTTAG 1260
Db	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATATAATTTGATAGATTTAG 1260
Qy	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAAAAATTTTACTATTTTG 1320
Db	1261	AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAAAAATTTTACTATTTTG 1320
Qy	1321	AGAAAGTTTGAATGTATAGAAATTTGAGTCGGACATCCATAAAGGAAAGTGTATAGAT 1380
Db	1321	AGAAAGTTTGAATGTATAGAAATTTGAGTCGGACATCCATAAAGGAAAGTGTATAGAT 1380
Qy	1381	TAAATGGGACAGAGGAGTAAATACCTTTATGATATATAAAATTTTGTATTTTCAATTTCA 1440
Db	1381	TAAATGGGACAGAGGAGTAAATACCTTTATGATATATAAAATTTTGTATTTTCAATTTCA 1440
Qy	1441	TAAAGATTATAAATCTATGTTTATAATGATAATAAATTTTAAAAATAATATATATTAAT 1500
Db	1441	TAAAGATTATAAATCTATGTTTATAATGATAATAAATTTTAAAAATAATATATATTAAT 1500
Qy	1501	CTGATTAGTCGATACCGCTTTTATAATTTTACAACTAGAGTAAATGAAATAAATCAG 1560
Db	1501	CTGATTAGTCGATACCGCTTTTATAATTTTACAACTAGAGTAAATGAAATAAATCAG 1560
Qy	1561	TTATCTGAAAAGCAAAATAATATCTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAGTTCATG 1620
Db	1561	TTATCTGAAAAGCAAAATAATATCTTTGTAAAAACAGCGTTCGGTCAAAATGGGAAGTTCATG 1620
Qy	1621	TGTAATCAATAGTTTATAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA 1680
Db	1621	TGTAATCAATAGTTTATAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA 1680
Qy	1681	AAATTAATAAATTAATTTAGAGCATGGAGTTCACGGGACATTTGAGAGCAGCTAGACT 1740
Db	1681	AAATTAATAAATTAATTTAGAGCATGGAGTTCACGGGACATTTGAGAGCAGCTAGACT 1740
Qy	1741	GTTTGAACAATGTATGTCGGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATTAAT 1800

Db 1741 GTTTGAACAATGATGTCCGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Qy 1801 GCATTCTAG 1809
Db 1801 GCATTCTAG 1809

RESULT 4

E40090 2048 bp DNA linear PAT 31-JAN-2002
LOCUS Plant promoter and terminator.
DEFINITION E40090
ACCESSION E40090
VERSION E40090.1 GI:18627206
KEYWORDS JP 2000166577-A/4.
SOURCE unidentifed
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 2048)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 4 20-JUN-2000;
SUMITOMO CHEM CO LTD

COMMENT OS Daucus carota L.
PN JP 2000166577-A/4
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//C12N5/10,C12R1:91), PC
C12N15/00,
CC C12N5/00, (C12N5/00, C12R1:91)

FEATURES

source
FH Key Location/Qualifiers
FT promoter Location/Qualifiers
1..2048
(1)..(2048).
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 85.2%; Score 1748; DB 6; Length 2048;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2048; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 CATGTGTGCCCTACAGACATAGGCGCTGTTGGTTGAGAGACGAGCGTCTCTGA 60
Db |||||||
Qy 1 CATGTGTGCCCTACAGACATAGGCGCTGTTGGTTGAGAGACGAGCGTCTCTGA 60
Db |||||||
Qy 61 CTCTCTCTCTTTTGACCTGTTGTATAAAGAGTAGAAATATTTTAAAAAGCTGCGAA 120
Db 61 CTCTCTCTCTTTTGACCTGTTGTATAAAGAGTAGAAATATTTTAAAAAGCTGCGAA 120
Qy 121 TACTAACTTCTCTCACAACCTCGGCTCTTTTCCAAAACCTTATTAACCTTTTACT 180
Db |||||||
Qy 121 TACTAACTTCTCTCACAACCTCGGCTCTTTTCCAAAACCTTATTAACCTTTTACT 180
Db |||||||
Qy 181 TCTCATTTCTACTCCACTCTTTGTATAGCAAGAAATCACTCTTTTAAAGCTAACCCA 240
Db |||||||
Qy 181 TCTCATTTCTACTCCACTCTTTGTATAGCAAGAAATCACTCTTTTAAAGCTAACCCA 240
Db |||||||
Qy 241 AACGGCCTCAATAAAAGATCAATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
Db |||||||
Qy 241 AACGGCCTCAATAAAGATCAATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT 300
Db |||||||
Qy 301 GAACAGGGTTATTTTAACTGTCAACAAATCTTAATAATTTTAACTGGCCGGTGAACA 360
Db |||||||
Qy 301 GAACAGGGTTATTTTAACTGTCAACAAATCTTAATAATTTTAACTGGCCGGTGAACA 360
Db |||||||
Qy 361 CCGTCTTCCAAGATAATATTTTAACTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Db |||||||
Qy 361 CCGTCTTCCAAGATAATATTTTAACTGTAGCTCCCTTTTAAACCAATTC----GC 416

Qy 421 AGGACGACTTAGGTGAATACACATTTGTACTGTGAGTCTTTAAACAAAGAAACAAGTGGTTC 480
Db |||||||
Qy 417 AGGACGACTTAGGTGAATACACATTTGTACTGTGAGTCTTTAAACAAAGAAACAAGTGGTTC 476
Db |||||||
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGCTACTATACTTT 540
Db |||||||
Qy 477 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGCTACTATACTTT 536
Db |||||||
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATAATATGCTCCATCCCAAGGATAAGT 600
Db |||||||
Qy 537 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATAATATGCTCCATCCCAAGGATAAGT 596
Db |||||||
Qy 601 AAAATCCCGTTTAAACCAAGTTTGTATATATATATTTTACACTTACAAAGAGATATTCGT 660
Db |||||||
Qy 597 AAAATCCCGTTTAAACCAAGTTTGTATATATATATTTTACACTTACAAAGAGATATTCGT 656
Db |||||||
Qy 661 AATACTTTTACACGACGAGAGACTTAGGTCAAAATGGAACGCTGGTAAACAGCCTAGACT 720
Db |||||||
Qy 657 AATACTTTTACACGACGAGAGACTTAGGTCAAAATGGAACGCTGGTAAACAGCCTAGACT 716
Db |||||||
Qy 721 TGGTCACCTGATATAATAGATTAATTTGTAGTATAATATAGTAGGATCTCAATGACATATAA 780
Db |||||||
Qy 717 TGGTCACCTGATATAATAGATTAATTTGTAGTATAATATAGTAGGATCTCAATGACATATAA 776
Db |||||||
Qy 781 ATTAGAGCTATTAAATTAAGTTTACTATAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Db |||||||
Qy 777 ATTAGAGCTATTAAATTAAGTTTACTATAATAAAGAGAGGTTAGTAAACAGAAAGCAGGTA 836
Db |||||||
Qy 841 AAAACAAGAGCTTCTGCTGTGTTGTTAGTTGTTGAGGCTCATTTCTTTAAAAGTAATG 900
Db |||||||
Qy 837 AAAACAAGAGCTTCTGCTGTGTTGTTAGTTGTTGAGGCTCATTTCTTTAAAAGTAATG 896
Db |||||||
Qy 901 TAAACTCATCTAAAGACACATAGAAATTTAGTACAGGTTTAAACCTTTTACAAAGATTTATA 960
Db |||||||
Qy 897 TAAACTCATCTAAAGACACATAGAAATTTAGTACAGGTTTAAACCTTTTACAAAGATTTATA 956
Db |||||||
Qy 961 TTAACGAAAATCAATTTTATAACAATGTCTCTCGGCTGTCAATTAATAAGGATCACTTAC 1020
Db |||||||
Qy 957 TTAACGAAAATCAATTTTATAACAATGTCTCTCGGCTGTCAATTAATAAGGATCACTTAC 1016
Db |||||||
Qy 1021 TGATCATCCATTAACAACTTGTTHAAAACAATTCATAGAGATAAAATATCTTACAATGAA 1080
Db |||||||
Qy 1017 TGATCATCCATTAACAACTTGTTHAAAACAATTCATAGAGATAAAATATCTTACAATGAA 1076
Db |||||||
Qy 1081 AAGAAGGACAATGTCTCTTTGAAAAAACAATAGTACTCCCTCGCTCCCTCTGAAAGT 1140
Db |||||||
Qy 1077 AAGAAGGACAATGTCTCTTTGAAAAAACAATAGTACTCCCTCGCTCCCTCTGAAAGT 1136
Db |||||||
Qy 1141 ATACATATGGATTGGACACGAGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Db |||||||
Qy 1137 ATACATATGGATTGGACACGAGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1196
Db |||||||
Qy 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGGGGACCCACCAATATATATTAATGTAGATTTAG 1260
Db |||||||
Qy 1197 AAAGAGAAAGAAAGTGGGTAAAGTAGGGGACCCACCAATATATATTAATGTAGATTTAG 1256
Db |||||||
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTTTATATTTATATAAAAATTTTACTATTTTG 1320
Db |||||||
Qy 1257 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTTTATATTTATATAAAAATTTTACTATTTTG 1316
Db |||||||
Qy 1321 AGAAAGTTTTGAAATGTATAGAAATTTAGTGGGACATCCATAAAAAGGAAAGTGTATAGAA 1380
Db |||||||
Qy 1317 AGAAAGTTTTGAAATGTATAGAAATTTAGTGGGACATCCATAAAAAGGAAAGTGTATAGAA 1376
Db |||||||
Qy 1381 TAAATGGGACAGAGGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGAATTTCA 1440
Db |||||||
Qy 1377 TAAATGGGACAGAGGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGAATTTCA 1436
Db |||||||
Qy 1441 TAAGATTATAAATCTATGTTTATATATATATATATATTTTAAAAATAATACTATATTAAT 1500
Db |||||||
Qy 1437 TAAGATTATAAATCTATGTTTATATATATATATATATTTTAAAAATAATACTATATTAAT 1496
Db |||||||
Qy 1501 CTGATTAGTCGATTACCGCCTTTTATAAATTTTACAATACTGAGTAATATGAATAAATCAG 1560


```

|||||
1497 CTGATTAGTCGATTACCGCCTTTTATAAATTTTAACTACTGAGTAATATGAATAAATCAG 1556
Qy |||||||
1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACACAGCGTTCCGGTCAAAATGGGAAATTCATG 1620
Db |||||||
1557 TTATCTGAAAGCAAAATAATATCTTTGTAAACACAGCGTTCCGGTCAAAATGGGAAATTCATG 1616
Qy |||||||
1621 TGTATTCAAATAGTTTTTAATAATAAAGTAAATTTTAAATTAATTTGTTTAAATTTTGTTCAGA 1680
Db |||||||
1617 TGTATTCAAATAGTTTTTAATAATAAAGTAAATTTTAAATTAATTTGTTTAAATTTTGTTCAGA 1676
Qy |||||||
1681 AATTTAAATAAATAATTTAGCATGCGGAGTTTCAAGGCGATCTATGAGCAGCAGTACTAGACT 1740
Db |||||||
1677 AATTTAAATAAATAATTTAGCATGCGGAGTTTCAAGGCGATCTATGAGCAGCAGTACTAGACT 1736
Qy |||||||
1741 GTTTGAACAATGTATGTCGCGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db |||||||
1737 GTTTGAACAATGTATGTCGCGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy |||||||
1801 GCATCTAGATAATACATCTTTTCAAAATTTCAAAACAACAGCTTTAACTTTTCTTTTCAACG 1860
Db |||||||
1797 GCATCTAGATAATACATCTTTTCAAAATTTCAAAACAACAGCTTTAACTTTTCTTTTCAACG 1856
Qy |||||||
1861 GATTGGAATCCTTTTCTAAACTTTTAAATTAATAAATAAATAAATGCAATTTTGAATATTTATC 1920
Db |||||||
1857 GATTGGAATCCTTTTCTAAACTTTTAAATAAATAAATAAATAAATGCAATTTTGAATATTTATC 1916
Qy |||||||
1921 AACACCTCAACATTTAGTGTAGCGTACTATATAATAGGTGCTCTTGGTGTCTACTATCAT 1980
Db |||||||
1917 AACACCTCAACATTTAGTGTAGCGTACTATATAATAGGTGCTCTTGGTGTCTACTATCAT 1976
Qy |||||||
1981 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
Db |||||||
1977 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2036
Qy |||||||
2041 ATTCTAATATC 2052
Db |||||||
2037 ATTCTAATATC 2048

RESULT 5
E40089
LOCUS E40089 2048 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40089
VERSION E40089.1 GI:18627205
KEYWORDS JP 2000166577-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
1 (bases 1 to 2048)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 3 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT
OS Daucus carota L.
PN JP 2000166577-A/3
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC
C12N15/00,
PC C12N5/00, (C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(2048).
1..2048
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
ORIGIN

```

```

Query Match      84.5%; Score 1733; DB 6; Length 2048;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2033; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 CATGTGTGCCTACACACATAGGGCTGTTTGGTTGAGAGAAGCAGCAAGCTGCTCTGA 60
Db |||||||
Qy 1 CATGTGTGCCTACACACATAGGGCTGTTTGGTTGAGAGAAGCAGCAAGCTGCTCTGA 60
Db |||||||
Qy 61 CTTCCTCTCTCTTGTGACCTGTTTGTATATAAGAAAGTAGAAATATTTTAAAAAGCTCGAA 120
Db |||||||
Qy 61 CTTCCTCTCTCTTGTGACCTGTTTGTATATAAGAAAGTAGAAATATTTTAAAAAGCTCGAA 120
Db |||||||
Qy 121 TACTAACTCTCTCTCACAACTTCCGCTTCTTTCCAAACACTTTATTAATCTTTTACT 180
Db |||||||
Qy 121 TACTAACTCTCTCTCACAACTTCCGCTTCTTTCCAAACACTTTATTAATCTTTTACT 180
Db |||||||
Qy 181 TCTCATTTCTACCTCCACTTCTTTGCTATAGCAAGAAATCCTTCTTTAAGCTTAACCCA 240
Db |||||||
Qy 181 TCTCATTTCTACCTCCACTTCTTTGCTATAGCAAGAAATCCTTCTTTAAGCTTAACCCA 240
Db |||||||
Qy 241 AACGGCTCTCAATAAAAAGATCATTTCAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
Db |||||||
Qy 241 AACGGCTCTCAATAAAAAGATCATTTCAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
Db |||||||
Qy 301 GAACAGGTTATTTTAAAGCTGTCAACAAATCTAATAATTTTAACTCTGCGCGGTGAACA 360
Db |||||||
Qy 301 GAACAGGTTATTTTAAAGCTGTCAACAAATCTAATAATTTTAACTCTGCGCGGTGAACA 360
Db |||||||
Qy 361 CCGTCTTCCAAGATAATATATTTTAAATTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Db |||||||
Qy 361 CCGTCTTCCAAGATAATATATTTTAAATTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Db |||||||
Qy 421 AGGACGACTTGTAGTGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGACAAGTGGTTC 480
Db |||||||
Qy 421 AGGACGACTTGTAGTGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGACAAGTGGTTC 480
Db |||||||
Qy 481 ATGCTCAGCCATCAAAATTCAGAAAACCCGACACACACTCTATCCAGTACTATCTTT 540
Db |||||||
Qy 481 ATGCTCAGCCATCAAAATTCAGAAAACCCGACACACACTCTATCCAGTACTATCTTT 540
Db |||||||
Qy 541 TGCCGGAATGCTCTCAAAATGTTTTTATATGTATAAATAATGCCCATCAAGGATAAGT 600
Db |||||||
Qy 541 TGCCGGAATGCTCTCAAAATGTTTTTATGTATAAATAATGCCCATCAAGGATAAGT 600
Db |||||||
Qy 601 AAAATTCCTGTTTAAACAGTTTGTATATATATATGTTTACACTTTACAGAGGATATTCGT 660
Db |||||||
Qy 601 AAAATTCCTGTTTAAACAGTTTGTATATATATATGTTTACACTTTACAGAGGATATTCGT 660
Db |||||||
Qy 661 AATACTTTTACAGCAGCAAGAGACTTAGGTCAAAAATGGAGCGCTGGTAAACAGCCTAGACT 720
Db |||||||
Qy 661 AATACTTTTACAGCAGCAAGAGACTTAGGTCAAAAATGGAGCGCTGGTAAACAGCCTAGACT 720
Db |||||||
Qy 721 TGGTCACTGATAAATAGATAATTTGTAGTATAATATAGTAGGATCTACAATGACATTTAAA 780
Db |||||||
Qy 721 TGGTCACTGATAAATAGATAATTTGTAGTATAATATAGTAGGATCTACAATGACATTTAAA 780
Db |||||||
Qy 781 ATTAGAGCTAATTAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAACAGGTA 840
Db |||||||
Qy 781 ATTAGAGCTAATTAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAACAGGTA 840
Db |||||||
Qy 841 AAAACAAGAGCTTGTCTGCTGTTTGTAGTTTGTGAGCTCATTTCTTTAAAAGTAAATG 900
Db |||||||
Qy 841 AAAACAAGAGCTTGTCTGCTGTTTGTAGTTTGTGAGCTCATTTCTTTAAAAGTAAATG 896
Db |||||||
Qy 901 TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAGCTTTTAAAGAAATTTATA 960
Db |||||||
Qy 997 TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAGCTTTTAAAGAAATTTATA 956
Db |||||||
Qy 961 TTAACGAAAATCATTTTATAACATGTCTCTCGGCTGCTCATTTATTAATAGGATCACTTAC 1020
Db |||||||
Qy 957 TTAACGAAAATCATTTTATAACATGTCTCTCGGCTGCTCATTTATTAATAGGATCACTTAC 1016
Db |||||||

```

Qy 1021 TGATCATCCATTAAACCTTGTGTTAAAAACAATTCATGAGATAAAATATCTTCAATGAA 1080
Db 1017 TGATCATCCATTAAACCTTGTGTTAAAAACAATTCATGAGATAAAATATCTTCAATGAA 1076
Qy 1081 AAGAGGACATGCTCTTTGAAAAACAATAGTACTCCCTCCGCTCTGAAATGT 1140
Db 1077 AAGAGGACATGCTCTTTGAAAAACAATAGTACTCCCTCCGCTCTGAAATGT 1136
Qy 1141 ATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Db 1137 ATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1196
Qy 1201 AAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTAG 1260
Db 1197 AAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTAG 1256
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAATTTACTATTG 1320
Db 1257 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAAATTTACTATTG 1316
Qy 1321 AGAAAGTTTTGAAATGTATAGAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
Db 1317 AGAAAGTTTTGAAATGTATAGAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1376
Qy 1381 TAAATGGGACAGAGGGAGTAAATACCTTTATGATATATAAATTTTGTATTGATTCA 1440
Db 1377 TAAATGGGACAGAGGGAGTAAATACCTTTATGATATATAAATTTTGTATTGATTCA 1436
Qy 1441 TAAGATTATAAATCTATGTTATATGATATATAAATTTTAAAAATATATCTATATTAAT 1500
Db 1437 TAAGATTATAAATCTATGTTATATGATATATAAATTTTAAAAATATATCTATATTAAT 1496
Qy 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTCAATACCTGAGTAATATCAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCTTTTATAATTTTCAATACCTGAGTAATATCAATAAATCAG 1556
Qy 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAATGGGAAGTTCATG 1620
Db 1557 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAATGGGAAGTTCATG 1616
Qy 1621 TGATTCAATAGTTTTAATAAAGTAATTTTAAATTTAATTTGTTATTTTTGTTTCAGA 1680
Db 1617 TGATTCAATAGTTTTAATAAAGTAATTTTAAATTTAATTTGTTATTTTTGTTTCAGA 1676
Qy 1681 AATTTAAATATAATTTAGTACGAGGAGTTCACGGGCATCATTTAGCAGACCTAGACT 1740
Db 1677 AATTTAAATATAATTTAGTACGAGGAGTTCACGGGCATCATTTAGCAGACCTAGACT 1736
Qy 1741 GTTTGAAACAATGTATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTGAAACAATGTATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATTTCTAGATAATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACG 1860
Db 1797 GCATTTCTAGATAATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACG 1856
Qy 1861 GATTGGAATCTTTTCAAACTTTTAAAAATAAATAAATGCAATTTATGTAATATTATC 1920
Db 1857 GATTGGAATCTTTTCAAACTTTTAAAAATAAATAAATGCAATTTATGTAATATTATC 1916
Qy 1921 AACACCTCAACATGTATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
Db 1917 AACACCTCAACATGTATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1976
Qy 1981 CACATCAATCTTACACCAAACTTGAGCTTAATTTTCTACTATTCTCAGCAAT 2037
Db 1977 CACATCAATCTTACACCAAACTTGAGCTTAATTTTCTACTATTCTCAGCAAT 2033

RESULT 6
ARI146856
LOCUS
DEFINITION Sequence 6 from patent US 6218598. linear PAT 08-AUG-2001

ACCESSION ARI146856
VERSION ARI146856.1 GI:15110045
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 140)
AUTHORS Ishige,F., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: US 6218598-A 6 17-APR-2001;
FEATURES Location/Qualifiers
 source
 1..140
 /organism="unknown"
 /mol_type="unassigned DNA"
ORIGIN
Query Match 2.8%; Score 57; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACATCAATCTT 1992
Db 23 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACATCAATCTT 79

RESULT 7
LOCUS ARI146857/C
DEFINITION Sequence 7 from patent US 6218598. 140 bp DNA linear PAT 08-AUG-2001
ACCESSION ARI146857
VERSION ARI146857.1 GI:15110046
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 140)
AUTHORS Ishige,F., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: US 6218598-A 7 17-APR-2001;
FEATURES Location/Qualifiers
 source
 1..140
 /organism="unknown"
 /mol_type="unassigned DNA"
ORIGIN
Query Match 2.8%; Score 57; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACATCAATCTT 1992
Db 122 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACATCAATCTT 66

RESULT 8
LOCUS E55069
DEFINITION Plant promoter.
ACCESSION E55069
VERSION E55069.1 GI:18625255
KEYWORDS JP 2000083679-A/6.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 140)
AUTHORS Ishige,I., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: JP 2000083679-A 6 28-MAR-2000;
COMMENT SUMITOMO CHEM CO LTD
OS Artificial Sequence
FN JP 2000083679-A/6
PD 28-MAR-2000
PF 12-JUL-1999 JP 1999197240

```
PR      IKUHARU ISHIGE, SATOMI NISHIKAWA, KENJI OBDA
PI      C12N15/09,A01H5/00,C12N1/21,C12N5/10//((C12N15/09,C12R1:91), PC
PC      (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91), PC
(C12N5/00,C12R1:91)
CC      Key Location/Qualifiers
FH      1..140
FT      /organism='Artificial Sequence'.
FEATURES
source   1..140 Location/Qualifiers
         /organism="synthetic construct"
         /mol_type="genomic DNA"
         /db_xref="taxon:32630"
ORIGIN
Query Match      2.8%; Score 57; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 1992
         |||||
Db      23 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 79
         |||||
RESULT 9
E55070/c
LOCUS      140 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter.
ACCESSION E55070
VERSION E55070.1 GI:18625256
KEYWORDS JP 2000083679-A/7.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 140)
AUTHORS Ishige,I., Nishikawa,S. and Oeda,K.
TITLES Plant promoter
JOURNAL Patent: JP 2000083679-A 7 28-MAR-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Artificial Sequence
PN JP 2000083679-A/7
PD 28-MAR-2000
PF 12-JUL-1999 JP 1999197240
PR
PI IKUHARU ISHIGE, SATOMI NISHIKAWA, KENJI OBDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//((C12N15/09,C12R1:91), PC
(C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91), PC
(C12N5/00,C12R1:91)
CC
FH      Key Location/Qualifiers
FT      1..140
FT      /organism='Artificial Sequence'.
FEATURES
source   1..140 Location/Qualifiers
         /organism="synthetic construct"
         /mol_type="genomic DNA"
         /db_xref="taxon:32630"
ORIGIN
Query Match      2.8%; Score 57; DB 6; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 1992
         |||||
Db      122 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 66
         |||||
RESULT 10
AR146852
LOCUS      246 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 2 from patent US 6218598.
```

```
AR146852
VERSION AR146852.1 GI:15110041
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,F., Nishikawa,S. and Oeda,K.
TITLES Plant promoter
JOURNAL Patent: US 6218598-A 2 17-APR-2001;
FEATURES Location/Qualifiers
source   1..246
         /organism="unknown"
         /mol_type="unassigned DNA"
ORIGIN
Query Match      2.8%; Score 57; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 1992
         |||||
Db      130 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 186
         |||||
RESULT 11
E55065
LOCUS      246 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter.
ACCESSION E55065
VERSION E55065.1 GI:18625251
KEYWORDS JP 2000083679-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,I., Nishikawa,S. and Oeda,K.
TITLES Plant promoter
JOURNAL Patent: JP 2000083679-A 2 28-MAR-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000083679-A/2
PD 28-MAR-2000
PF 12-JUL-1999 JP 1999197240
PR
PI IKUHARU ISHIGE, SATOMI NISHIKAWA, KENJI OBDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//((C12N15/09,C12R1:91), PC
(C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91), PC
(C12N5/00,C12R1:91)
CC
FH      Key Location/Qualifiers
FT      1..246
FT      /organism='Daucus carota L.'.
FEATURES
source   1..246 Location/Qualifiers
         /organism="unidentified"
         /mol_type="genomic DNA"
         /db_xref="taxon:32644"
ORIGIN
Query Match      2.8%; Score 57; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 1992
         |||||
Db      130 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCATCAATCTT 186
         |||||
RESULT 12
AR076816
LOCUS      247 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 1 from patent US 5959176.
```

```
ACCESSION AR076816
VERSION AR076816.1 GI:10003562
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 247)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 1 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..247
/mol_type="unassigned DNA"
ORIGIN
Query Match 2.8%; Score 57; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1992
Db 131 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 187
RESULT 13
E15124
LOCUS E15124 247 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15124
VERSION E15124.1 GI:5709807
KEYWORDS JP 1998052273-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 1 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/1
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..247
FT /organism="Daucus carota L."
FT /clone="pCR16G1-Xb"
FT promoter 1<..<247.
FT Location/Qualifiers
FEATURES source 1..247
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 2.8%; Score 57; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1992
Db 131 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 187
RESULT 14
```

```
AR076817
LOCUS AR076817 2042 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959176.
ACCESSION AR076817
VERSION AR076817.1 GI:10003563
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2042)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 2 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..2042
/mol_type="unassigned DNA"
ORIGIN
Query Match 2.8%; Score 57; DB 6; Length 2042;
Best Local Similarity 100.0%; Pred. No. 8.2e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1992
Db 1926 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1982
RESULT 15
E15125
LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/2
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..2042
FT /organism="Daucus carota L."
FT promoter 1<..<2042.
FT Location/Qualifiers
FEATURES source 1..2042
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 2.8%; Score 57; DB 6; Length 2042;
Best Local Similarity 100.0%; Pred. No. 8.2e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1992
Db 1926 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1982
```

Search completed: October 28, 2005, 03:19:37
Job time : 9012 secs

RESULT 1
AAA37959
ID AAA37959 standard; DNA; 2052 BP.
XX

Db 361 CCGTCTCCAGATAATATATATTTTAAATTTTGTAGCCCTCCCTTTTAAACCAATTCGCATGC 420
Qy 421 AGGACGACTTAGTGGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGAAACAAGTGGTTC 480
Db 421 AGGACGACTTAGTGGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGAAACAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAAACACTCTATCAGCTACTATATCTTT 540
Db 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAAACACTCTATCAGCTACTATATCTTT 540
Qy 541 TGGCCGAATGCTCTCAAAATGTTTTTATATGTATGTAATAATGCCCCATCCAGGATTAAGT 600
Db 541 TGGCCGAATGCTCTCAAAATGTTTTTATATGTATGTAATAATGCCCCATCCAGGATTAAGT 600
Qy 601 AAAATCCCGTTTAAACAGTTTGTATATATATATGTTTACACTTTACAAGAGATATTCGT 660
Db 601 AAAATCCCGTTTAAACAGTTTGTATATATATATGTTTACACTTTACAAGAGATATTCGT 660
Qy 661 AATACTTTTACAGCACAAGAGACTTAGGTCAAAAATGGAGCGCTGGTAAACAGCCTAGACT 720
Db 661 AATACTTTTACAGCACAAGAGACTTAGGTCAAAAATGGAGCGCTGGTAAACAGCCTAGACT 720
Qy 721 TGGTCACTGATAATAGATAATGTTAGTATATATAGTAGGATCTACAATGACATTTAAA 780
Db 721 TGGTCACTGATAATAGATAATGTTAGTATATATAGTAGGATCTACAATGACATTTAAA 780
Qy 781 ATTAGACTATTAATTAAGTTACTTAATAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Db 781 ATTAGACTATTAATTAAGTTACTTAATAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
Qy 841 AAAACAGAGCTTCTCTGCTGTGTTTGTAGTGTGAGTCTATTTTAAAAGTAATG 900
Db 841 AAAACAGAGCTTCTCTGCTGTGTTTGTAGTGTGAGTCTATTTTAAAAGTAATG 900
Qy 901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA 960
Db 901 TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA 960
Qy 961 TTTAAACGAAAATCAATTTTAAACATGTCTCTCGCTGTCTATATAATAGGGATCACCTTAC 1020
Db 961 TTTAAACGAAAATCAATTTTAAACATGTCTCTCGCTGTCTATATAATAGGGATCACCTTAC 1020
Qy 1021 TGATCATCTTAATAAACCTTTGTTAAACAAATTCATGAGATAAATAATCTTCAATAGAA 1080
Db 1021 TGATCATCTTAATAAACCTTTGTTAAACAAATTCATGAGATAAATAATCTTCAATAGAA 1080
Qy 1081 AAGAAAGACATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGT 1140
Db 1081 AAGAAAGACATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGT 1140
Qy 1141 ATACATATGGATTGGACACCGAGACTAAGAAAAATGTATATAAGTAATGTAGAGTAAAAAG 1200
Db 1141 ATACATATGGATTGGACACCGAGACTAAGAAAAATGTATATAAGTAATGTAGAGTAAAAAG 1200
Qy 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCACCAATATATATTCATAGATTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCACCAATATATATTCATAGATTAG 1260
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTG 1320
Qy 1321 AGAAAGTTTTGAAATGTATAGAAATTTAGTGGGACATCCATAAAAGGAAAGTGTATAGAA 1380
Db 1321 AGAAAGTTTTGAAATGTATAGAAATTTAGTGGGACATCCATAAAAGGAAAGTGTATAGAA 1380
Qy 1381 TAAATGGGACAGGGAGTAACTTTATGATATATAAATTTTGTATTTTGTATTTCA 1440
Db 1381 TAAATGGGACAGGGAGTAACTTTATGATATATAAATTTTGTATTTTGTATTTCA 1440
Qy 1441 TAAGATTATAAATCTATGTTATATATGATAATAAATTTTAAATAATACTATATTAAT 1500
Db 1441 TAAGATTATAAATCTATGTTATATATGATAATAAATTTTAAATAATACTATATTAAT 1500

RESULT 2

AAA37961

ID AAA37961 standard; DNA; 2052 BP.

XX

AC AAA37961;

XX

DT 18-AUG-2000 (first entry)

XX

DE Carrot promoter sequence #2.

XX

KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

XX

OS Daucus carota.

XX

PN WO200020613-A1.

XX

PD 13-APR-2000.

XX

PF 28-SEP-1999; 99WO-JP005303.

XX

PR 02-OCT-1998; 98JP-00281124.

XX

PA (SUMO) SUMITOMO CHEM CO LTD.

XX

PI Nishikawa S, Oeda K;

XX

DR WPI; 2000-303791/26.

XX

PT New plant promoters and terminators from Daucus carota L., useful in
plant breeding, for e.g. controlling fertilities of plants.

XX

Claim 1; Page 78-79; 81pp; English.

This sequence represents a carrot promoter. The invention relates to plant promoters and terminators from *Daucus carota* L. which are capable of expressing a gene of interest in plants. The invention also includes a chimeric gene characterized in that it comprises the promoter and a desired gene linked to each other in the form capable of functioning. A method of producing a transformant comprises introducing the promoter, the chimeric gene or a vector comprising the promoter and a desired gene or terminator sequence into a host cell. The plant promoters and terminators are useful in plant breeding, for e.g. fertilities of plants may be controlled by expressing, in the host cells, a sense or antisense gene of a male sterility related gene such as S-locus-specific RNase gene

Query Match		99.3%	Score 2037	DB 3	Length 2052
Best Local Similarity		100.0%	Pred. No. 0		
Matches 2037	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	1	CATGTGTGCCCTACAGCACATAGGGCCCTGTTGGTTGAGAGAACAGAGCTGCTTCGA	60		
Db	1	CATGTGTGCCCTACAGCACATAGGGCCCTGTTGGTTGAGAGAACAGAGCTGCTTCGA	60		
Qy	61	CTTCTTCTCTTTTGACCTGTTGTATAAGAAAGTAGAAATATTTTAAAAAGCTGCGAA	120		
Db	61	CTTCTTCTCTTTTGACCTGTTGTATAAGAAAGTAGAAATATTTTAAAAAGCTGCGAA	120		
Qy	121	TACTAACTTCTCTCTCAACACTTCGGCTCTTTTCCAAACACTTTATTAACCTTTTACT	180		
Db	121	TACTAACTTCTCTCTCAACACTTCGGCTCTTTTCCAAACACTTTATTAACCTTTTACT	180		
Qy	181	TCTCATTTTCTACTCCACTTCTTTGCTATAGCAAGAAATCACTTCTTTAAGCTAAACCA	240		
Db	181	TCTCATTTTCTACTCCACTTCTTTGCTATAGCAAGAAATCACTTCTTTAAGCTAAACCA	240		
Qy	241	AACGGCTCAATAAAGATCATTAATGTATCTTTCAATTTTAGGATTAACATAGT	300		
Db	241	AACGGCTCAATAAAGATCATTAATGTATCTTTCAATTTTAGGATTAACATAGT	300		
Qy	301	GAACAGGGTATTTTAAACGGTCAACAAATCTTAATAATTTTACCTGGCGGTGAACA	360		
Db	301	GAACAGGGTATTTTAAACGGTCAACAAATCTTAATAATTTTACCTGGCGGTGAACA	360		
Qy	361	CCGCTTTCCAAGATAATATATTTTAAATTTGTAGCTCCCTTTTAAACAAATTCGATGC	420		
Db	361	CCGCTTTCCAAGATAATATATTTTAAATTTGTAGCTCCCTTTTAAACAAATTCGATGC	420		
Qy	421	AGGACGACTTAGGTGAATACACATGTACTGTGAGTCTTTTAAACAAAGAAACAGTGGTC	480		
Db	421	AGGACGACTTAGGTGAATACACATGTACTGTGAGTCTTTTAAACAAAGAAACAGTGGTC	480		
Qy	481	ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACTCTATCCAGTACTATACTTT	540		
Db	481	ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACTCTATCCAGTACTATACTTT	540		
Qy	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATATATATATATATATATATATATATAT	600		
Db	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATATATATATATATATATATATATATAT	600		
Qy	601	AAAATTCCTGGTTAAACAGTTTGTATATATATATATATATATATATATATATATAT	660		
Db	601	AAAATTCCTGGTTAAACAGTTTGTATATATATATATATATATATATATATATATAT	660		
Qy	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT	720		
Db	661	AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT	720		
Qy	721	TGGTCACGTATAAATAGATAATTTGTAGTATAATATAGTAGGATCTCAAAAGCATTA	780		
Db	721	TGGTCACGTATAAATAGATAATTTGTAGTATAATATAGTAGGATCTCAAAAGCATTA	780		
Qy	781	ATTAGACTATTAATTAAGTTTACTAATAAATAAGAGAGAGTTTAGTAAACAGAAAGCAGTA	840		

```
Db 1861 GATTGGAATCCTTTCTAAACTTTTAAATAAAAAATGCAATATTGTAATATTATC 1920
Qy 1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
Db 1921 AACACCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCAT 1980
Qy 1981 CACATCAATCTTACACACACAAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAAT 2037
Db 1981 CACATCAATCTTACACACACAAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAAT 2037

RESULT 3
AAA37964
ID AAA37964 standard; DNA; 2056 BP.
XX
AC AAA37964;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO20020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
  plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
  introducing a mutation into a carrot promoter. The invention relates to
  plant promoters and terminators from Daucus carota L. which are capable
  of expressing a gene of interest in plants. The invention also includes a
  chimeric gene characterized in that it comprises the promoter and a
  desired gene linked to each other in the form capable of functioning. A
  method of producing a transformant comprises introducing the promoter,
  the chimeric gene or a vector comprising the promoter and a desired gene
  or terminator sequence into a host cell. The plant promoters and
  terminators are useful in plant breeding, for e.g. fertilities of plants
  may be controlled by expressing, in the host cells, a sense or antisense
  gene of a male sterility related gene such as S-locus-specific RNase gene
  Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;

Query Match 88.2%; Score 1809; DB 3; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAGAAGCTGCTCTGA 60
Db 1 CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAAGCAGAAGCTGCTCTGA 60
Qy 61 CTTCCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATATTTTAAAGCTGCGAA 120
Db 61 CTTCCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATATTTTAAAGCTGCGAA 120
Qy 121 TACTAACTCTCTCACAACTCCGCTCTCTTTTCCAAACACCTTATTAATTTTACT 180
  |||
```

```
Db 121 TACTAACTCTCTCACAACTCCGCTCTCTTTTCCAAACACCTTATTAATTTTACT 180
Qy 181 TCTCATTTTCTACTCCACTCTCTTTTGCTATAGCAAGAAATCACCTTCTTTTAAGCTAACCCA 240
Db 181 TCTCATTTTCTACTCCACTCTCTTTTGCTATAGCAAGAAATCACCTTCTTTTAAGCTAACCCA 240
Qy 241 AACGGCTCAATAAAGATCATTAATATGATCTTTCAATTTTAGGATAACAATAGT 300
Db 241 AACGGCTCAATAAAGATCATTAATATGATCTTTCAATTTTAGGATAACAATAGT 300
Qy 301 GAACAGGTTATTTTAAACGTGCAACAAATCTAATAATTTTACCTGCGCGTGAACA 360
Db 301 GAACAGGTTATTTTAAACGTGCAACAAATCTAATAATTTTACCTGCGCGTGAACA 360
Qy 361 CCGTCTTTCAAAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAAATTCGCATGC 420
Db 361 CCGTCTTTCAAAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAAATTCGCATGC 420
Qy 421 AGGACGACTTAGGTGAATACACATTGTACTGTAGTCTTTTAAACAAGAAACAAGTGGTTC 480
Db 421 AGGACGACTTAGGTGAATACACATTGTACTGTAGTCTTTTAAACAAGAAACAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTCACAAACCCGACACACACTCTATCCACGTACTATCTTT 540
Db 481 ATGCTCAGCCATCAAAATTCACAAACCCGACACACACTCTATCCACGTACTATCTTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAAGGATAAGT 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAAGGATAAGT 600
Qy 601 AAAATCCCGTTTAAACGAGTTGTTAATATATATATATGTATTTTACATCAAGAGGATATTCGT 660
Db 601 AAAATCCCGTTTAAACGAGTTGTTAATATATATATATGTATTTTACATCAAGAGGATATTCGT 660
Qy 661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAAATGGACGCTGGTAAACAGCCCTAGACT 720
Db 661 AATACTTTTAGACGACAAGAGACTTAGGTCAAAAAATGGACGCTGGTAAACAGCCCTAGACT 720
Qy 721 TGGTCACTGATAAATAGATAAATTTGTAGTATAATATAGTAGGATCTCAATATGACATTTAAA 780
Db 721 TGGTCACTGATAAATAGATAAATTTGTAGTATAATATAGTAGGATCTCAATATGACATTTAAA 780
Qy 781 ATTAGAGCTATTAAATTAAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAACAGGTA 840
Db 781 ATTAGAGCTATTAAATTAAGTTACTTAATAATAGAGAGGTTAGTAAACAGAAACAGGTA 840
Qy 841 AAAACAAGAGCTTGTGCTGTGTGTTAGTGTGTTGTGAGCTCATTTCTTTAAAAGTAAATG 900
Db 841 AAAACAAGAGCTTGTGCTGTGTGTTAGTGTGTTGTGAGCTCATTTCTTTAAAAGTAAATG 900
Qy 901 TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
Db 901 TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGAATTTATA 960
Qy 961 TTAACGAAATCATTTTATAACATGTCTCTCGGTGTCATTTAATAGGATCAGCTTAC 1020
Db 961 TTAACGAAATCATTTTATAACATGTCTCTCGGTGTCATTTAATAGGATCAGCTTAC 1020
Qy 1021 TGATCATCCATTAACACCTTTGTTAAACAAATCAATGAGATAAAATATCTTACAATGAA 1080
Db 1021 TGATCATCCATTAACACCTTTGTTAAACAAATCAATGAGATAAAATATCTTACAATGAA 1080
Qy 1081 AAGAAGGACAAATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGT 1140
Db 1081 AAGAAGGACAAATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGT 1140
Qy 1141 ATACATATGATTTGGACACGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Db 1141 ATACATATGATTTGGACACGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAG 1200
Qy 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCACCAATATATAATTTGATAGATTTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCACCAATATATAATTTGATAGATTTAG 1260
```

1261 AAAAGTAGTTGAAAGTAGTGGTGGGATTTTATATATATAAAATTTACTATTTTG 1320
 1261 AAAAGTAGTTGAAAGTAGTGGTGGGATTTTATATATATAAAATTTACTATTTTG 1320
 1321 AGAAAGTTTGAATATGATAGATTTGAGTGGGATCCATATAAGGAAAGTGTATAGAA 1380
 1321 AGAAAGTTTGAATATGATAGATTTGAGTGGGATCCATATAAGGAAAGTGTATAGAA 1380
 1381 TAAATGGACAGAGGGAGTAACTTTATGATATATAAAATTTTGTATTTGATTTC 1440
 1381 TAAATGGACAGAGGGAGTAACTTTATGATATATAAAATTTTGTATTTGATTTC 1440
 1441 TAAGATTATAAATCTATGTTATATATGATATATAAAATTTTAAATATACTATATTAAT 1500
 1441 TAAGATTATAAATCTATGTTATATATGATATATAAAATTTTAAATATACTATATTAAT 1500
 1501 CTGATTAGTCGATTTACCGCTTTTATATATTAATTTTCAATCTGAGTAATGATAAATCAG 1560
 1501 CTGATTAGTCGATTTACCGCTTTTATATATTAATTTTCAATCTGAGTAATGATAAATCAG 1560
 1561 TTATCTGAAAGCAAAATATATCTTTGTAAGACAGCGTTCGGTCAAAATGGGAGTTTCATG 1620
 1561 TTATCTGAAAGCAAAATATATCTTTGTAAGACAGCGTTCGGTCAAAATGGGAGTTTCATG 1620
 1621 TGTATTCGAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTTGTATTTTTCAGA 1680
 1621 TGTATTCGAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTTGTATTTTTCAGA 1680
 1681 AATTTAAATATAATTTATGAGCATGGGAAGTTTCAGCGGCATCATGTAGAGCAGCTAGACT 1740
 1681 AATTTAAATATAATTTATGAGCATGGGAAGTTTCAGCGGCATCATGTAGAGCAGCTAGACT 1740
 1741 GTTTGAACAATGTATGTCGGGTGATCATCTATGACCTTCAACTCAACTAGTAGTAAT 1800
 1741 GTTTGAACAATGTATGTCGGGTGATCATCTATGACCTTCAACTCAACTAGTAGTAAT 1800
 1801 GCATTCTAG 1809
 1801 GCATTCTAG 1809
 RESULT 4
 AAA37962
 ID AAA37962 standard; DNA; 2048 BP.
 AC AAA37962;
 XX 18-AUG-2000 (first entry)
 XX Plasmid #1 DNA sequence used in mutation of promoter sequence.
 XX Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
 XX Daucus carota.
 XX WO200020613-A1.
 XX 13-APR-2000.
 XX 28-SEP-1999; 99WO-JP005303.
 XX 02-OCT-1998; 98JP-00281124.
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX Nishikawa S, Oeda K;
 XX WPI; 2000-303791/26.
 XX New Plant promoters and terminators from Daucus carota L., useful in
 XX plant breeding, for e.g. controlling fertilities of plants.

Example 8; Page 71-73; 81pp; English.

PS This sequence represents a plasmid sequence used in a method for
 XX introducing a mutation into a carrot promoter. The invention relates to
 CC plant promoters and terminators from Daucus carota L. which are capable
 CC of expressing a gene of interest in plants. The invention also includes a
 CC chimeric gene characterized in that it comprises the promoter and a
 CC desired gene linked to each other in the form capable of functioning. A
 CC method of producing a transformant comprises introducing the promoter,
 CC the chimeric gene or a vector comprising the promoter and a desired gene
 CC or terminator sequence into a host cell. The plant promoters and
 CC terminators are useful in plant breeding, for e.g. fertilities of plants
 CC may be controlled by expressing, in the host cells, a sense or antisense
 CC gene of a male sterility related gene such as S-locus-specific RNase gene
 XX
 SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 84.5%; Score 1733; DB 3; Length 2048;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2033; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 Qy 1 CATGTGTGCCCTACAGCAGATAGGGCTGTTGGTTGAGAGAGCAGAGCTGTTCTGA 60
 Db 1 CATGTGTGCCCTACAGCAGATAGGGCTGTTGGTTGAGAGAGCAGAGCTGTTCTGA 60
 Qy 61 CTTCCTCTCTCTCTGACCTGTTGTATATAAGAGAGTGAATATTTTAAAAAGCTGCGAA 120
 Db 61 CTTCCTCTCTCTCTGACCTGTTGTATATAAGAGAGTGAATATTTTAAAAAGCTGCGAA 120
 Qy 121 TACTAACTCTCTCTCAAACTTCGGCTCTCTTTTCCAAACACTTTATTAATTTTTTACT 180
 Db 121 TACTAACTCTCTCTCAAACTTCGGCTCTCTTTTCCAAACACTTTATTAATTTTTTACT 180
 Qy 181 TCTCATTTCTCTCTCACTCTTTGCTATAGAGAGAGTAATCACTTCTTTAAGCTAAACCCA 240
 Db 181 TCTCATTTCTCTCTCACTCTTTGCTATAGAGAGAGTAATCACTTCTTTAAGCTAAACCCA 240
 Qy 241 AACGGCTCAATAAAGATCATTTCAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
 Db 241 AACGGCTCAATAAAGATCATTTCAATAATGTATCTTTCAATTTTAGGATAACAATACGT 300
 Qy 301 GAACAGGGTTATTTTAAACGTGTCAACAAATCTAATAATTTTAACTGGCCGGTGAACA 360
 Db 301 GAACAGGGTTATTTTAAACGTGTCAACAAATCTAATAATTTTAACTGGCCGGTGAACA 360
 Qy 361 CCGTCTCCAGAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCATGC 420
 Db 361 CCGTCTCCAGAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCATGC 420
 Qy 421 AGGACGACTTTAGGTGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGACAAGTGGTTC 480
 Db 421 AGGACGACTTTAGGTGAATACACATTTGTACTGTGAGTCTTTTAAACAAAGACAAGTGGTTC 480
 Qy 481 ATGCTCAGCCATCAAAATTCGACAAACCCGACACACACTCTATCCACGACTATATCTTT 540
 Db 481 ATGCTCAGCCATCAAAATTCGACAAACCCGACACACACTCTATCCACGACTATATCTTT 540
 Qy 541 TGGCCGAGATCTCTCAAAATGTTTTTATATGTAAATATGCCCCATCCAGGATAAGT 600
 Db 541 TGGCCGAGATCTCTCAAAATGTTTTTATATGTAAATATGCCCCATCCAGGATAAGT 600
 Qy 601 AAAATTTCCCGTTTAAACCAAGTTTGTATATATATATTTTACACTTTACAAGAGGATATTCGT 660
 Db 601 AAAATTTCCCGTTTAAACCAAGTTTGTATATATATATTTTACACTTTACAAGAGGATATTCGT 660
 Qy 661 AATACCTTTTACGACGACAGACTTAGGTCAAAATAGGACGCTGGTGTAAACAGCCTAGACT 720
 Db 661 AATACCTTTTACGACGACAGACTTAGGTCAAAATAGGACGCTGGTGTAAACAGCCTAGACT 720
 Qy 721 TGGTCACTGATAAATAGATAATTTGTAGTATATATATAGTAGGATCTCAATGACATTAAA 780
 Db 721 TGGTCACTGATAAATAGATAATTTGTAGTATATATATAGTAGGATCTCAATGACATTAAA 780

XX 18-OCT-2002; 2002JP-00304115.
XX
XX 18-OCT-2002; 2002JP-00304115.
XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX
XX WPI; 2004-360986/34.
XX
XX Expression inducing promoter, useful for expressing foreign gene,
XX comprising first DNA for detecting transcription start point and having
XX minimum promoter function, linked to second DNA having expression
XX inducing promoter function.
XX
XX Disclosure; SEQ ID NO 4; 61pp; Japanese.
XX
XX The invention relates to an expression inducing promoter comprising a
XX first DNA linked to a second DNA at its 5' terminus, where the first DNA
XX has a region which determines a transcription start point of RNA
XX polymerase II and has minimum promoter function, and the second DNA has
XX expression inducing promoter function in a plant cell. The expression
XX inducing promoter is useful for expressing a foreign gene. This sequence
XX represents carrot DNA used in the scope of the invention.
XX
XX Sequence 196 BP; 66 A; 40 C; 19 G; 71 T; 0 U; 0 Other;
XX
XX Query Match 9.0%; Score 185; DB 12; Length 196;
XX Best Local Similarity 100.0%; Pred. No. 8.5e-74;
XX Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1853 TTTCAACGGATTGGAAATCCCTTTTCTAACTTTTAAATATAAAAAATGCAATTATTCTAA 1912
Db 1 TTTCAACGGATTGGAAATCCCTTTTCTAACTTTTAAATATAAAAAATGCAATTATTCTAA 60
XX
Qy 1913 TATTATCAACACCTCAACATTGATGTTAGCGTACTATATAAGTGCTCTTTGGTGCTCT 1972
Db 61 TATTATCAACACCTCAACATTGATGTTAGCGTACTATATAAGTGCTCTTTGGTGCTCT 120
XX
Qy 1973 ACTATCATCATCAATCTTTACACACCAACCTTTGAGCTTAATTTTCTACTTATTCTCA 2032
Db 121 ACTATCATCATCAATCTTTACACACCAACCTTTGAGCTTAATTTTCTACTTATTCTCA 180
XX
Qy 2033 GCAAT 2037
Db 181 GCAAT 185
XX
RESULT 7
AAZ49616/C
ID AAZ49616 standard; DNA; 140 BP.
XX
XX AC AAZ49616;
XX
XX 07-APR-2000 (first entry)
XX
XX Oligonucleotide-4 for synthesis of CR16.1 fragment for plant promoter.
XX
XX Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
XX soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
XX male sterility-related gene; ss.
XX
XX Daucus carota.
XX OS Synthetic.
XX
XX EP976832-A2.
XX
XX 02-FEB-2000.
XX
XX 13-JUL-1999; 99EP-00113732.
XX
XX 15-JUL-1998; 98JP-00200372.
XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX
XX Ishige F, Nishikawa S, Oeda K;
XX
XX WPI; 2000-128374/12.
XX
XX Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX
XX Disclosure; Page 13; 24pp; English.
XX
XX The present sequence is an oligonucleotide (+ chain) used to prepare a
XX CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX construction of a plant promoter. The promoter is used for controlling
XX the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX desaturase and S-locus type specific RNase gene (male sterility-related
XX gene) in a host cell especially a microorganism or a plant cell. The
XX transformed plant cells can be used to produce transgenic plants. The
XX promoter is compact and therefore suitable for higher expression of a
XX desired gene in a particular tissue compared to other host tissues
XX
XX Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;
XX
XX Query Match 2.8%; Score 57; DB 3; Length 140;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-15;
XX Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX

XX
PI Ishige F, Nishikawa S, Oeda K;
XX
XX DR WPI; 2000-128374/12.
XX
XX PT Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX
XX PS Disclosure; Page 14; 24pp; English.
XX
XX The present sequence is an oligonucleotide (- chain) used to prepare a
XX CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX construction of a plant promoter. The promoter is used for controlling
XX the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX desaturase and S-locus type specific RNase gene (male sterility-related
XX gene) in a host cell especially a microorganism or a plant cell. The
XX transformed plant cells can be used to produce transgenic plants. The
XX promoter is compact and therefore suitable for higher expression of a
XX desired gene in a particular tissue compared to other host tissues
XX
XX Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;
XX
XX Query Match 2.8%; Score 57; DB 3; Length 140;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-15;
XX Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1936 ATGTTAGCGTACTATATAAGTGCTCTTTGGTGCTCTTACTATCATCATCAATCTTT 1992
Db 122 ATGTTAGCGTACTATATAAGTGCTCTTTGGTGCTCTTACTATCATCATCAATCTTT 66
XX
RESULT 8
AAZ49615
ID AAZ49615 standard; DNA; 140 BP.
XX
XX AC AAZ49615;
XX
XX 07-APR-2000 (first entry)
XX
XX Oligonucleotide-3 for synthesis of CR16.1 fragment for plant promoter.
XX
XX Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
XX soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
XX male sterility-related gene; ss.
XX
XX Daucus carota.
XX OS Synthetic.
XX
XX EP976832-A2.
XX
XX 02-FEB-2000.
XX
XX 13-JUL-1999; 99EP-00113732.
XX
XX 15-JUL-1998; 98JP-00200372.
XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX
XX Ishige F, Nishikawa S, Oeda K;
XX
XX WPI; 2000-128374/12.
XX
XX Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX
XX PS Disclosure; Page 13; 24pp; English.
XX
XX The present sequence is an oligonucleotide (+ chain) used to prepare a
XX CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX construction of a plant promoter. The promoter is used for controlling
XX the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX

```
CC desaturase and S-locus type specific RNase gene (male sterility-related
CC gene) in a host cell especially a microorganism or a plant cell. The
CC transformed plant cells can be used to produce transgenic plants. The
CC promoter is compact and therefore suitable for higher expression of a
CC desired gene in a particular tissue compared to other host tissues
XX
SQ Sequence 140 BP; 44 A; 34 C; 17 G; 45 T; 0 U; 0 Other;
Query Match 2.8%; Score 57; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 2.1e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1936 ATGTTAGCGTACTATAAATAGTCTCTTGGTCTCTACTATCATCATCAATCTT 1992
Db 23 ATGTTAGCGTACTATAAATAGTCTCTTGGTCTCTACTATCATCATCAATCTT 79
RESULT 9
AAZ49611
ID AAZ49611 standard; DNA; 246 BP.
XX
AC AAZ49611;
XX
DT 07-APR-2000 (first entry)
XX
DE Carrot CR16.3 fragment for synthesis of plant promoter.
XX
KW Synthetic DNA; plant promoter; CR16.3 fragment; carrot; transgenic plant;
KW soybean glycinin; stearyl-ACP-desaturase gene;
KW male sterility-related gene; ds.
XX
OS Daucus carota.
XX
PN EP976832-A2.
XX
PD 02-FEB-2000.
XX
PF 13-JUL-1999; 99EP-00113732.
XX
PR 15-JUL-1998; 98JP-00200372.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Ishige F, Nishikawa S, Oeda K;
XX
WPI; 2000-128374/12.
XX
Novel promoter used to produce transgenic plants with higher expression
of a desired gene.
PS Claim 1; Page 11-12; 24pp; English.
XX
The present sequence is a CR16.3 fragment from carrot genomic DNA. It is
used for synthesis of a plant promoter which comprises nucleotides 112-
246, 54-246, or 1-246 of this sequence and a synthetic DNA. The promoter
is used for controlling the expression of a desired gene e.g. soybean
glycinin, stearyl-ACP-desaturase and S-locus type specific RNase gene
(male sterility-related gene) in a host cell especially a microorganism
or a plant cell. The transformed plant cells can be used to produce
transgenic plants. The promoter is compact and therefore suitable for
higher expression of a desired gene in a particular tissue compared to
other host tissues
XX
SQ Sequence 246 BP; 88 A; 50 C; 21 G; 87 T; 0 U; 0 Other;
Query Match 2.8%; Score 57; DB 3; Length 246;
Best Local Similarity 100.0%; Pred. No. 2e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1936 ATGTTAGCGTACTATAAATAGTCTCTTGGTCTCTACTATCATCATCAATCTT 1992
Db 130 ATGTTAGCGTACTATAAATAGTCTCTTGGTCTCTACTATCATCATCAATCTT 186
```

```
RESULT 10
AAV15143
ID AAV15143 standard; DNA; 247 BP.
XX
AC AAV15143;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX
WPI; 1998-122310/12.
XX
New carrot root gene, promoter and terminator - useful in genetic
engineering for directing root-specific gene expression.
PS Claim 1; Page 14; 3lpp; English.
XX
The present sequence represents a novel promoter, and is isolated from
the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
direct root-specific expression in plant cells. Since the promoter
enables expression of a desired protein in the roots of a plant, it is
useful in combat against pathogenic soil fungi and pests which are
difficult to kill by chemicals. It can also be used to improve the
nutritive value of edible root plants
XX
SQ Sequence 247 BP; 88 A; 50 C; 21 G; 88 T; 0 U; 0 Other;
Query Match 2.8%; Score 57; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1936 ATGTTAGCGTACTATAAATAGTCTCTTGGTCTCTACTATCATCATCAATCTT 1992
Db 131 ATGTTAGCGTACTATAAATAGTCTCTTGGTCTCTACTATCATCATCAATCTT 187
RESULT 11
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX
AC AAV15144;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
```



```

PD 18-FEB-1998.
XX
XX 12-AUG-1997; 97EP-00113923.
XX
XX 12-AUG-1996; 96JP-00212680.
XX
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX Torikai S, Oeda K;
XX
XX WPI; 1998-122310/12.
XX
XX New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
XX Claim 2; Page 15-16; 31pp; English.
XX
XX The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
XX Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
SQ
Query Match 2.8%; Score 57; DB 2; Length 2042;
Best Local Similarity 100.0%; Pred. No. 1.9e-15;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGGTCCTCTTGGTGTCTACTATCATCATCAATCTTT 1992
Db 1926 ATGTTAGCGTACTATAAATAGGTCCTCTTGGTGTCTACTATCATCATCAATCTTT 1982

Search completed: October 28, 2005, 05:33:54
Job time : 1122 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2005, 21:12:09 ; Search time 6850 Seconds
(without alignments)
11402.610 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcaataacattctaaatc 2052

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 50

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

- Database : EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_hic:*
 - 4: gb_est3:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_gss1:*
 - 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

No matches found

Search completed: October 28, 2005, 05:13:53
Job time : 6850 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 00:49:34 ; Search time 354 Seconds
(without alignments)
9484.858 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgcccctacagcaca.....gcaatacattctaaatc 2052

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 50

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description
Result No.	Score	Query Match	Length	ID	
c	1	57	2.8	140	3 US-09-352-608-6
	2	57	2.8	140	3 US-09-352-608-7
	3	57	2.8	246	3 US-09-352-608-2
	4	57	2.8	247	2 US-08-911-434A-1
	5	57	2.8	2042	2 US-08-911-434A-2

ALIGNMENTS

RESULT 1
US-09-352-608-6
; Sequence 6, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6

; LENGTH: 140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-6

Query Match 2.8%; Score 57; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 2.4e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1992
Db 23 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 79

RESULT 2
US-09-352-608-7/c
; Sequence 7, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-7

Query Match 2.8%; Score 57; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 2.4e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 1992
Db 122 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTTACTATCATCATCAATCTT 66

RESULT 3
US-09-352-608-2
; Sequence 2, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-352-608-2

Query Match 2.8%; Score 57; DB 3; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;

```
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTTACTATCATCATCAATCTT 1992
Db 130 ATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTTACTATCATCATCAATCTT 186

RESULT 4
US-08-911-434A-1
; Sequence 1, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 1:
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..247
US-08-911-434A-1

Query Match 2.8%; Score 57; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTTACTATCATCATCAATCTT 1992
Db 131 ATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTTACTATCATCATCAATCTT 187

RESULT 5
US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..2042
US-08-911-434A-2

Query Match 2.8%; Score 57; DB 2; Length 2042;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1936 ATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTTACTATCATCATCAATCTT 1992
Db 1926 ATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTTACTATCATCATCAATCTT 1982

Search completed: October 28, 2005, 07:54:11
Job time : 354 secs
```

Job time : 2333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 05:14:00 ; Search time 2333 Seconds
(without alignments)
7264.230 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtccctacagcaca.....gcaataacattctaataatc 2052

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9784742 seqs, 4129495052 residues

Word size : 50

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09E_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description

No matches found

Search completed: October 28, 2005, 10:22:19

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 00:21:14 ; Search time 8050 Seconds
(without alignment)
10406.073 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtcctacagcaca.....gcataacattctaataatc 2052

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 45554873 seqs, 20411521753 residues

Word size : 50

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*

61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*

62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*

63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*

64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*

65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*

66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*

67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*

68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*

69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*

70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*

71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*

72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*

73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*

74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*

75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*

76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*

77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*

78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*

79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*

80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*

81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*

93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*

94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*

100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*

101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*

103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*

104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*

105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*

106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*

107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*

108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*

109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*

110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*

111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*

113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*

114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*

115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*

116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2052	100.0	2052	35	US-09-806-197-1 Sequence 1, Appli
2	2037	99.3	2052	35	US-09-806-197-7 Sequence 7, Appli
3	1809	88.2	2056	35	US-09-806-197-5 Sequence 5, Appli
4	1733	84.5	2048	35	US-09-806-197-3 Sequence 3, Appli
5	1733	84.5	2048	35	US-09-806-197-4 Sequence 4, Appli
6	57	2.8	247	20	US-09-300-487-1 Sequence 1, Appli
7	57	2.8	2042	20	US-09-300-487-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
 US-09-806-197-1
 ; Sequence 1, Application US/09806197
 ; GENERAL INFORMATION:
 ; APPLICANT: NISHIKAWA, SATOMI
 ; APPLICANT: OEDA, KENJI
 ; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
 ; FILE REFERENCE: 7372-70911
 ; CURRENT APPLICATION NUMBER: US/09/806.197
 ; CURRENT FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 2052
 ; TYPE: DNA
 ; ORGANISM: Daucus carota
 US-09-806-197-1

Query Match 100.0%; Score 2052; DB 35; Length 2052;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CATGTGTGCCCTACAGCATATAGGCGCTGTTGGTTGAGAGAAGCAGAACTGCTTCTGA	60
Db	1	CATGTGTGCCCTACAGCATATAGGCGCTGTTGGTTGAGAGAAGCAGAACTGCTTCTGA	60
Qy	61	CTTCTCTCTCTTTTGACCTGTTGTATAAGAAAGTAGAATAATTTTAAAGCTGGAA	120
Db	61	CTTCTCTCTCTTTTGACCTGTTGTATAAGAAAGTAGAATAATTTTAAAGCTGGAA	120
Qy	121	TACTAACTTCTCTCACAACCTCGGCTCTTTTCCAAACACTTATTAACTTTTACT	180
Db	121	TACTAACTTCTCTCACAACCTCGGCTCTTTTCCAAACACTTATTAACTTTTACT	180
Qy	181	TCTCATTTCTACTCCACTTCTTGTCTATAAGCAAGAAATCACTCTTTTAAAGCTAACCA	240
Db	181	TCTCATTTCTACTCCACTTCTTGTCTATAAGCAAGAAATCACTCTTTTAAAGCTAACCA	240

Qy	241	AACGGCTCAATAAAGATCAATCATATAATGTATCTTTCAATTTTAGGATAACAATACGT	300
Db	241	AACGGCTCAATAAAGATCAATCATATAATGTATCTTTCAATTTTAGGATAACAATACGT	300
Qy	301	GAACAGGGTTATTTTAAACGTGTCAACAAATTTCTAATAATTTTACCTGCGCGGTGAACA	360
Db	301	GAACAGGGTTATTTTAAACGTGTCAACAAATTTCTAATAATTTTACCTGCGCGGTGAACA	360
Qy	361	CCGTCTTCCAAGATAATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGATGC	420
Db	361	CCGTCTTCCAAGATAATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGATGC	420
Qy	421	AGGACGACTTGTAGTGAATACACATTTGTCTGAGTCTTTTAAACAAAGACAGTGGTTC	480
Db	421	AGGACGACTTGTAGTGAATACACATTTGTCTGAGTCTTTTAAACAAAGACAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAAATTCACAAAACCCGACACACACTCTATCCACGTAATATCTTT	540
Db	481	ATGCTCAGCCATCAAAAATTCACAAAACCCGACACACACTCTATCCACGTAATATCTTT	540
Qy	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAAGATTAAGT	600
Db	541	TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAAGATTAAGT	600
Qy	601	AAATTTCCCGTTTAAACAGTTTGTAAATATATATATATATATATATATATATATATAT	660
Db	601	AAATTTCCCGTTTAAACAGTTTGTAAATATATATATATATATATATATATATATATAT	660
Qy	661	AATACTTTTACAGCACAAGAGACTTTAGGTCAAAAATGGAGCGTGGTAAACAGCCCTAGCT	720
Db	661	AATACTTTTACAGCACAAGAGACTTTAGGTCAAAAATGGAGCGTGGTAAACAGCCCTAGCT	720
Qy	721	TGCTCACTGATAAATAGATAAATTTGTAGTATAATATATAGTAGGATCTACAATGAATAA	780
Db	721	TGCTCACTGATAAATAGATAAATTTGTAGTATAATATATAGTAGGATCTACAATGAATAA	780
Qy	781	ATTAGAGCTAATAATTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAAACAGGTA	840
Db	781	ATTAGAGCTAATAATTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAAACAGGTA	840
Qy	841	AAAACAGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	841	AAAACAGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Qy	901	TAAACTGATCTAAAGCACAATAGAAAATTTAGTACAGGTTTAAACCTTTTACAAGAATTTATA	960
Db	901	TAAACTGATCTAAAGCACAATAGAAAATTTAGTACAGGTTTAAACCTTTTACAAGAATTTATA	960
Qy	961	TTAAACGAAATCAATTTTATAACATGCTCTCGGCTGCTCATTATAATAGGATCACTTAC	1020
Db	961	TTAAACGAAATCAATTTTATAACATGCTCTCGGCTGCTCATTATAATAGGATCACTTAC	1020
Qy	1021	TGATCATCCATTTAAACCTTTGTTTAAACAAATTTCAATGAGATAAAATATCTTACAATGAA	1080
Db	1021	TGATCATCCATTTAAACCTTTGTTTAAACAAATTTCAATGAGATAAAATATCTTACAATGAA	1080
Qy	1081	AAGAAGACAATGCTCTTTTGAATAAATAAATAAGTACTCCCTCCGCTCCCTCTGAAATGT	1140
Db	1081	AAGAAGACAATGCTCTTTTGAATAAATAAATAAGTACTCCCTCCGCTCCCTCTGAAATGT	1140
Qy	1141	ATACATATGGATTGGACACGAGACTAAGAAAATGTATAAAGTAATGTAGAGTAAAGAAAG	1200
Db	1141	ATACATATGGATTGGACACGAGACTAAGAAAATGTATAAAGTAATGTAGAGTAAAGAAAG	1200
Qy	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTAGATTTAG	1260
Db	1201	AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTAGATTTAG	1260
Qy	1261	AAAAGTAGTGAAGTAGTGGGTGGGATTTTATATATATAAATAATTTACTATTTTG	1320
Db	1261	AAAAGTAGTGAAGTAGTGGGTGGGATTTTATATATATAAATAATTTACTATTTTG	1320
Qy	1321	AGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT	1380

Db	1321	AGAAAGTTTGGAAATGTATAGAAATTTAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT	1380
Qy	1381	TAAATGGGACAGAGGAGTAACTTTATGATATATAAAATTTTGTATTTTTCATTTCA	1440
Db	1381	TAAATGGGACAGAGGAGTAACTTTATGATATATAAAATTTTGTATTTTTCATTTCA	1440
Qy	1441	TAAGATTATAAATCTATGTTATATAATGATAATATAATTTTAAAAATAAATCTATATAAT	1500
Db	1441	TAAGATTATAAATCTATGTTATATAATGATAATATAATTTTAAAAATAAATCTATATAAT	1500
Qy	1501	CTGATTAGTCGATTACCGCCTTTTATATTTTCAATCTAGAGTAATATGATATAAATCAG	1560
Db	1501	CTGATTAGTCGATTACCGCCTTTTATATTTTCAATCTAGAGTAATATGATATAAATCAG	1560
Qy	1561	TTATCTGAAAGCAAAATAATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCATG	1620
Db	1561	TTATCTGAAAGCAAAATAATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCATG	1620
Qy	1621	TGTAATCAATAGTTTAAATAAAGTAAATTTTAAATTTAAATTTTGTATTTTGTTCAGA	1680
Db	1621	TGTAATCAATAGTTTAAATAAAGTAAATTTTAAATTTAAATTTTGTATTTTGTTCAGA	1680
Qy	1681	AATTTAAATTAATTTAGCATGGAAGTTTCAAGGCAATCATTTGAGCAGCACTAGACT	1740
Db	1681	AATTTAAATTAATTTAGCATGGAAGTTTCAAGGCAATCATTTGAGCAGCACTAGACT	1740
Qy	1741	GTATGAAATGATGTCGGGTGACATCTATGACCTTCAACTCAAACTAGTGAATAAT	1800
Db	1741	GTATGAAATGATGTCGGGTGACATCTATGACCTTCAACTCAAACTAGTGAATAAT	1800
Qy	1801	GCATTTAGATAATCTTTTCAAAATTTCAAAACACAGCTTTAACTTTTCTTTTCAACG	1860
Db	1801	GCATTTAGATAATCTTTTCAAAATTTCAAAACACAGCTTTAACTTTTCTTTTCAACG	1860
Qy	1861	GATTGGAATCTTTTCAAACTTTTAAATAAATAAATAAATGCAATATTTGTAATTTATC	1920
Db	1861	GATTGGAATCTTTTCAAACTTTTAAATAAATAAATAAATGCAATATTTGTAATTTATC	1920
Qy	1921	AACACCTCAACATTTAGTGTAGCTACTATATAATAGTGTCTTGGTCTCTACTATCAT	1980
Db	1921	AACACCTCAACATTTAGTGTAGCTACTATATAATAGTGTCTTGGTCTCTACTATCAT	1980
Qy	1981	CACATCAATCTTACACCAAAACCTTGAGCTTAAATTTTCTACTATTTCTCAGCAATAAC	2040
Db	1981	CACATCAATCTTACACCAAAACCTTGAGCTTAAATTTTCTACTATTTCTCAGCAATAAC	2040
Qy	2041	ATTCTAAATATC 2052	
Db	2041	ATTCTAAATATC 2052	

RESULT 2
US-09-806-197-7
; Sequence 7, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806.197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-806-197-7

Query Match 99.3%; Score 2037; DB 35; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2037; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CATGTGTGCGCTACAGCACATAGGCGCTGTTTGGTTCAGAGAAAGAGCTGCTTCTGA	60
Db	1	CATGTGTGCGCTACAGCACATAGGCGCTGTTTGGTTCAGAGAAAGAGCTGCTTCTGA	60
Qy	61	CTTCTCTCTTTTGACCTGTTTGTATATAAGAAAGTAGAAAATATTTTAAAAAGCTGCGAA	120
Db	61	CTTCTCTCTTTTGACCTGTTTGTATATAAGAAAGTAGAAAATATTTTAAAAAGCTGCGAA	120
Qy	121	TACTAACTCTCTCTCAAACTTCCGCTTCTTTTCCAAAACACTTTATTTAACTTTTACT	180
Db	121	TACTAACTCTCTCTCAAACTTCCGCTTCTTTTCCAAAACACTTTATTTAACTTTTACT	180
Qy	181	TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAAACCCA	240
Db	181	TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAAACCCA	240
Qy	241	AACGGCTCAATATAAGATCAATCAATAATGTATCTTTTCAATTTTAGATTAACAATACGT	300
Db	241	AACGGCTCAATATAAGATCAATCAATAATGTATCTTTTCAATTTTAGATTAACAATACGT	300
Qy	301	GAACAGGTTATTTTAAAGCTGTCACAAATTTCTAATAATTTTACCTGGCGGCTGAACA	360
Db	301	GAACAGGTTATTTTAAAGCTGTCACAAATTTCTAATAATTTTACCTGGCGGCTGAACA	360
Qy	361	CCGTCTTCCAAAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGCATGC	420
Db	361	CCGTCTTCCAAAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGCATGC	420
Qy	421	AGGACGCTTTAGGTGAATACACATTTGATCTTTTAAACCAAGAAAGAGTGGTTC	480
Db	421	AGGACGCTTTAGGTGAATACACATTTGATCTTTTAAACCAAGAAAGAGTGGTTC	480
Qy	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGCTACTATACTTT	540
Db	481	ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGCTACTATACTTT	540
Qy	541	TGGCCGAATGCTTCTCAAAATTTTATATATATATATATATATATATATATATATATAT	600
Db	541	TGGCCGAATGCTTCTCAAAATTTTATATATATATATATATATATATATATATATATAT	600
Qy	601	AAAATCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATAT	660
Db	601	AAAATCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATAT	660
Qy	661	AATACCTTTTACAGCAAGAGACTTAGGTCAAAAATGACACGCTTAAACAGCCCTAGACT	720
Db	661	AATACCTTTTACAGCAAGAGACTTAGGTCAAAAATGACACGCTTAAACAGCCCTAGACT	720
Qy	721	TGGTCACTGATAAATAGATAATTTGTATATATATATATATATATATATATATATATAT	780
Db	721	TGGTCACTGATAAATAGATAATTTGTATATATATATATATATATATATATATATATAT	780
Qy	781	ATTAGAGCTATTAATTAAGTTACTATAATAAGAGAGGTTAGTAAACAGAAACAGGTA	840
Db	781	ATTAGAGCTATTAATTAAGTTACTATAATAAGAGAGGTTAGTAAACAGAAACAGGTA	840
Qy	841	AAAACAGAGCTTGTCTGCTGTTTGTATTTGTGAGCTCATTTCTTTTAAAGTAAATG	900
Db	841	AAAACAGAGCTTGTCTGCTGTTTGTATTTGTGAGCTCATTTCTTTTAAAGTAAATG	900
Qy	901	TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA	960
Db	901	TAAACTGATCTAAAGCACATAGAAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA	960
Qy	961	TTAAACGAAATCATTTTATAACATGTCTCTCGGCTGCTCATTTAATAAGGATCACTTAC	1020
Db	961	TTAAACGAAATCATTTTATAACATGTCTCTCGGCTGCTCATTTAATAAGGATCACTTAC	1020
Qy	1021	TGATCATCCATTAACCTTTGTTAAACCAATTTCAATGAGATAAATAATCTTCAATAGAA	1080
Db	1021	TGATCATCCATTAACCTTTGTTAAACCAATTTCAATGAGATAAATAATCTTCAATAGAA	1080

1081 AAGAGGACATGCTCTTTGAAAAACAATAGGTACTCCCTCCGCTCCTGAAATGT 1140
 1081 AAGAGGACATGCTCTTTGAAAAACAATAGGTACTCCCTCCGCTCCTGAAATGT 1140
 1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
 1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
 1201 AAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAACATATATATATGATGATTTAG 1260
 1201 AAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAACATATATATATGATGATTTAG 1260
 1261 AAAAGTAGTTCAAGTAGTGGGTGGGTGGGTATTTATATATAAAAAATTTACTATTTTG 1320
 1261 AAAAGTAGTTCAAGTAGTGGGTGGGTGGGTATTTATATATAAAAAATTTACTATTTTG 1320
 1321 AGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAAGGAAAGTGTATAGAA 1380
 1321 AGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAAGGAAAGTGTATAGAA 1380
 1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTTGTATTTGATTTCA 1440
 1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTTGTATTTGATTTCA 1440
 1441 TAAGATTATAAATCTATGTTATATAGTAATATAAATTTTAAAAATATATATATTAAT 1500
 1441 TAAGATTATAAATCTATGTTATATAGTAATATAAATTTTAAAAATATATATATTAAT 1500
 1501 CTGATTAGTGCATTACCGCCCTTTTATAATTTTACAATCTGAGTAATATGAATAAATCAG 1560
 1501 CTGATTAGTGCATTACCGCCCTTTTATAATTTTACAATCTGAGTAATATGAATAAATCAG 1560
 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCA 1620
 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCA 1620
 1621 TGTATTCAATAGTTTAAATATAAAGTAATTTTAAATTAATTTTGTATTTTGTTCAGA 1680
 1621 TGTATTCAATAGTTTAAATATAAAGTAATTTTAAATTAATTTTGTATTTTGTTCAGA 1680
 1681 AATTTAAAAATAATTAATGAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACT 1740
 1681 AATTTAAAAATAATTAATGAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACT 1740
 1741 GTTTGAAACAATGTATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
 1741 GTTTGAAACAATGTATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
 1801 GCATTCTAGATACATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTCTTTTCAACG 1860
 1801 GCATTCTAGATACATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTCTTTTCAACG 1860
 1861 GATTGGAATCCCTTTTCAAACTTTTAAAAATAAAAAAATGCAATTTTGTAAATTTTATC 1920
 1861 GATTGGAATCCCTTTTCAAACTTTTAAAAATAAAAAAATGCAATTTTGTAAATTTTATC 1920
 1921 AACACCTCAACATTTGATGTTAGCGTATATATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
 1921 AACACCTCAACATTTGATGTTAGCGTATATATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
 1981 CACATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAAT 2037
 1981 CACATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTATTTCTCAGCAAT 2037

RESULT 3

US-09-806-197-5
 ; Sequence 5, Application US/09806197
 ; GENERAL INFORMATION:
 ; APPLICANT: NISHIKAWA, SATOMI
 ; APPLICANT: OEDA, KENJI
 ; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
 ; FILE REFERENCE: 7372-70911

Query Match 88.2%; Score 1809; DB 35; Length 2056;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ; CURRENT APPLICATION NUMBER: US/09/806,197
 ; CURRENT FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 2056
 ; TYPE: DNA
 ; ORGANISM: Daucus carota
 ; US-09-806-197-5

Qy 1 CATGTGTGCCTCAGACACATAGGGCTGTTGGTTGAGAGAAGCAGAACTGCTTCTGA 60
 Db 1 CATGTGTGCCTCAGACACATAGGGCTGTTGGTTGAGAGAAGCAGAACTGCTTCTGA 60
 Qy 61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAAGTAAAAATATTTTAAAAAGCTGCGAA 120
 Db 61 CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAAAGTAAAAATATTTTAAAAAGCTGCGAA 120
 Qy 121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACATTTATTAATCTTTTACT 180
 Db 121 TACTAACTTCTCTCTCACAACTTCCGCTTCTTTTCCAAACACATTTATTAATCTTTTACT 180
 Qy 181 TCTCAATTTCTCTCTCTCTTTGGCTATTAAGAAAGAAATCAGTCTCTTTTAAGCTAACCCA 240
 Db 181 TCTCAATTTCTCTCTCTCTTTGGCTATTAAGAAAGAAATCAGTCTCTTTTAAGCTAACCCA 240
 Qy 241 AAGCGGCTCAATAAAGAAAGTCAATAAATGATCTTTCAATTTTAGGATAACAATACGT 300
 Db 241 AAGCGGCTCAATAAAGAAAGTCAATAAATGATCTTTCAATTTTAGGATAACAATACGT 300
 Qy 301 GAACAGGGTTATTTTAAAGTGTCAACAAATTTCTAATAATTTTACCTGGCCGGTGAACA 360
 Db 301 GAACAGGGTTATTTTAAAGTGTCAACAAATTTCTAATAATTTTACCTGGCCGGTGAACA 360
 Qy 361 CGGTCTTCCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCGATGC 420
 Db 361 CGGTCTTCCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCGATGC 420
 Qy 421 AGGACGACTTAGGTGAATACACATTTGACCTGAGCTTCTTTTAAACAAAGAAAGTGGTTC 480
 Db 421 AGGACGACTTAGGTGAATACACATTTGACCTGAGCTTCTTTTAAACAAAGAAAGTGGTTC 480
 Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATCTTT 540
 Db 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATCTTT 540
 Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGT 600
 Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGT 600
 Qy 601 AATAATCCCGTTTAAACAGTTTGTATATATATGTTTACACTTTTACAGAGGATATTCGT 660
 Db 601 AATAATCCCGTTTAAACAGTTTGTATATATATGTTTACACTTTTACAGAGGATATTCGT 660
 Qy 661 AATACTTTTAGACGACAAAGACTTAGGTCAAAAATGGAGCTGGTAAACAGCCCTAGACT 720
 Db 661 AATACTTTTAGACGACAAAGACTTAGGTCAAAAATGGAGCTGGTAAACAGCCCTAGACT 720
 Qy 721 TGGTCACTGATAAATAGATAATTTGTTAGTATATATATAGTAGGATCTCAATAGACATTA 780
 Db 721 TGGTCACTGATAAATAGATAATTTGTTAGTATATATATAGTAGGATCTCAATAGACATTA 780
 Qy 781 ATTAGAGCTATTAATTAAGTTTACTAATAAATAGAGAGGTTAGTAACAGAAACAGGTA 840
 Db 781 ATTAGAGCTATTAATTAAGTTTACTAATAAATAGAGAGGTTAGTAACAGAAACAGGTA 840
 Qy 841 AAAACAAGAGCTTCTGCTGCTGTTGTTAGTTGTTGAGCTCATTTCTTTAAAAAGTAATG 900
 Db 841 AAAACAAGAGCTTCTGCTGCTGTTGTTAGTTGTTGAGCTCATTTCTTTAAAAAGTAATG 900

```

Qy 901 TAAACTGATCTAAAGCAGATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATA 960
Db 901 TAAACTGATCTAAAGCAGATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATA 960
Qy 961 TTAACGAAATCAATTTTATAACATGTCTCTCGGCTGTCAATATAAAGGATCACTTAC 1020
Db 961 TTAACGAAATCAATTTTATAACATGTCTCTCGGCTGTCAATATAAAGGATCACTTAC 1020
Qy 1021 TGATCATCCATTAACCTTTTAAACCAATTCATGAGTAAATATCTTACAAATGAA 1080
Db 1021 TGATCATCCATTAACCTTTTAAACCAATTCATGAGTAAATATCTTACAAATGAA 1080
Qy 1081 AAGAAGGACATGTCTCTTTGAAACCAATTCATGAGTAAATATCTTACAAATGAA 1140
Db 1081 AAGAAGGACATGTCTCTTTGAAACCAATTCATGAGTAAATATCTTACAAATGAA 1140
Qy 1141 ATACATATGAGTATGAGCAGCGAGACTAAGAAAAATGTATAAAGTAAATGTAGATGAT 1200
Db 1141 ATACATATGAGTATGAGCAGCGAGACTAAGAAAAATGTATAAAGTAAATGTAGATGAT 1200
Qy 1201 AAGAGAGAAAGAAAGTGGGTAAAGTACGCGGACCCACCAATATATAATGATAGATGAT 1260
Db 1201 AAGAGAGAAAGAAAGTGGGTAAAGTACGCGGACCCACCAATATATAATGATAGATGAT 1260
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAATTTTACTATTTTG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATAAATTTTACTATTTTG 1320
Qy 1321 AGAAAGTTTGAATATAGAAATGAGTGGGACATCCATAAAGGAAAGTATAGAAAT 1380
Db 1321 AGAAAGTTTGAATATAGAAATGAGTGGGACATCCATAAAGGAAAGTATAGAAAT 1380
Qy 1381 TAAATGGACAGAGGAGTAACTTTATGATATATAAATTTTGTATTTGATTTCA 1440
Db 1381 TAAATGGACAGAGGAGTAACTTTATGATATATAAATTTTGTATTTGATTTCA 1440
Qy 1441 TAAGATTATAAATCTATGTTATATATATATATAAATTTTAAATATACTATATTAAT 1500
Db 1441 TAAGATTATAAATCTATGTTATATATATATAAATTTTAAATATACTATATTAAT 1500
Qy 1501 CTGATTAGTCGATACCGCTTTTATAATTTTCAATCTGAGTAAATGAAATGAAATCAG 1560
Db 1501 CTGATTAGTCGATACCGCTTTTATAATTTTCAATCTGAGTAAATGAAATGAAATCAG 1560
Qy 1561 TTATCTGAAAGCAAAATATCTTTTAAACAGCGTTCGGTCAATGGAAGTTTCATG 1620
Db 1561 TTATCTGAAAGCAAAATATCTTTTAAACAGCGTTCGGTCAATGGAAGTTTCATG 1620
Qy 1621 TGATTCAATAGTTTAAATATAAAGTAAATTTTAAATTTTAAATTTTGTATTTTTCAGA 1680
Db 1621 TGATTCAATAGTTTAAATATAAAGTAAATTTTAAATTTTAAATTTTGTATTTTTCAGA 1680
Qy 1681 AATTTAAATATAATTTATGAGCATGGAAGTTTCAAGGATCATTTGAGCAGCATAGACT 1740
Db 1681 AATTTAAATATAATTTATGAGCATGGAAGTTTCAAGGATCATTTGAGCAGCATAGACT 1740
Qy 1741 GTTTGACAAATGATGTCGGGTGACATCTATGACCTTTCACTCAAACTAGTGAATTAAT 1800
Db 1741 GTTTGACAAATGATGTCGGGTGACATCTATGACCTTTCACTCAAACTAGTGAATTAAT 1800
Qy 1801 GCATTCTAG 1809
Db 1801 GCATTCTAG 1809

```

RESULT 4

```

US-09-806-197-3
; Sequence 3, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS

```

```

; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806,197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Daucus carota
; US-09-806-197-3

Query Match      84.5%; Score 1733; DB 35; Length 2048;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2033; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 CATGTGTGCGCTCAGACACATAGGGCCCTGTTGGTTCAGAGAGACAGAGCTGTTCTGA 60
Db 1 CATGTGTGCGCTCAGACACATAGGGCCCTGTTGGTTCAGAGAGACAGAGCTGTTCTGA 60
Qy 61 CTTCCTCTCTTTTGACCTGTTTGTATAAAGAGTAAATATTTTAAAGAGCTGCGAA 120
Db 61 CTTCCTCTCTTTTGACCTGTTTGTATAAAGAGTAAATATTTTAAAGAGCTGCGAA 120
Qy 121 TACTAACTTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTATTAACCTTTTACT 180
Db 121 TACTAACTTCTCTCTCAAACTTCCGCTTCTTTTCCAAACACATTTATTAACCTTTTACT 180
Qy 181 TCTCATTTCTACTCCACTCTTTGCTATAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240
Db 181 TCTCATTTCTACTCCACTCTTTGCTATAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240
Qy 241 AAGCGCTCAATATAAAGATCAATTAATATGATCTTTCAATTTTAGGATAACAATAGCT 300
Db 241 AAGCGCTCAATATAAAGATCAATTAATATGATCTTTCAATTTTAGGATAACAATAGCT 300
Qy 301 GAAACAGGGTTATTTTAAACGTTGTCACAAATTCATAAATTTTAACTGGCCGGTGAACA 360
Db 301 GAAACAGGGTTATTTTAAACGTTGTCACAAATTCATAAATTTTAACTGGCCGGTGAACA 360
Qy 361 CCGTCTTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Db 361 CCGTCTTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
Qy 421 AGACAGCTTTAGTGTGAATACACATTTGACGTGTCAGTCTTTTAAACAAAGACAGTGGTTC 480
Db 421 AGACAGCTTTAGTGTGAATACACATTTGACGTGTCAGTCTTTTAAACAAAGACAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTCGACAAACCCGACACACACCTCTATCCAGTACTATACCTTT 540
Db 481 ATGCTCAGCCATCAAAATTCGACAAACCCGACACACACCTCTATCCAGTACTATACCTTT 540
Qy 541 TGCGCGAATGCTTCTCAAAATGTTTATATATATATATATATATATATATATATATATATAT 600
Db 541 TGCGCGAATGCTTCTCAAAATGTTTATATATATATATATATATATATATATATATATATAT 600
Qy 601 AAAATTTCCCTTTAAACAGTTTGTATATATATATATATATATATATATATATATATATAT 660
Db 601 AAAATTTCCCTTTAAACAGTTTGTATATATATATATATATATATATATATATATATATAT 660
Qy 661 AATACTTTTAGACGACAGAGACTTAGGTCAAAATAGGACGCTGGTAAACAGCCTAGACT 720
Db 661 AATACTTTTAGACGACAGAGACTTAGGTCAAAATAGGACGCTGGTAAACAGCCTAGACT 720
Qy 721 TGCTCACTGATAAATAGATAAATTTGTAGTATATATATATATATATATATATATATATATAT 780
Db 721 TGCTCACTGATAAATAGATAAATTTGTAGTATATATATATATATATATATATATATATATAT 780
Qy 781 ATTAGAGCTAATTAATTAAGTTTACTTAATAATAGAGAGGTTAGTAAACAGAAACAGGTA 840
Db 781 ATTAGAGCTAATTAATTAAGTTTACTTAATAATAGAGAGGTTAGTAAACAGAAACAGGTA 840
Qy 841 AAAACAGAGCTTGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 841 AAAACAGAGCTTGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT

```


	D	b	657	AATACTTTTATAGACGACAAGAGACTTAGGTCAAAAATGGACGCCTGGTAAACAGCGCTAGACT	716
	Q	y	721	TGGTCACCTGATAAATAGATAAATCTGTAGTATAAATATAGTAGGATCTCAAAATGACATTAAA	780
	D	b	717	TGGTCACCTGATAAATAGATAAATCTGTAGTATAAATATAGTAGGATCTAACATGACATTAAA	776
	Q	y	781	ATTAGAGCTATTAAATTAAGTTTACTAAATAAAGAGAGGTTAGTAAACAGAGAAGCAGGTA	840
	D	b	777	ATTAGAGCTATTAAATTAAGTTTACTAAATAAATAGAGAGGTTAGTAAACAGAGAAGCAGGTA	836
	Q	y	841	AAAAAAGAGCGCTCGCTGCTGTCTTTAGTTTGTCGAGCTCATTTCTTTAAAGTAATG	900
	D	b	837	AAAAAAGAGCGCTCGCTGCTGTCTTTAGTTTGTCGAGCTCATTTCTTTAAAGTAATG	896
	Q	y	901	TAAACTGATCTTAAAGCACATAGAAAATTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA	960
	D	b	897	TAAACTGATCTTAAAGCACATAGAAAATTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA	956
	Q	y	961	TTAAACGAAATCAATTTTATAAACAATGCTCTCCGCTGCTCATTAATAAGGATCACCTTAC	1020
	D	b	957	TTAAACGAAATCAATTTTATAAACAATGCTCTCCGCTGCTCATTAATAAGGATCACCTTAC	1016
	Q	y	1021	TGATCATCCATTAAACCTTTGTTTAAACAACAATTCGAATGAGATAAAATATCTTACAAATGAA	1080
	D	b	1017	TGATCATCCATTAAACCTTTGTTTAAACAACAATTCGAATGAGATAAAATATCTTACAAATGAA	1076
	Q	y	1081	AAGAAGGACAATGCTCTCTTGGAAAAAACAATAGCTACTCCCTCCGCTCCCTCTGAAATGT	1140
	D	b	1077	AAGAAGGACAATGCTCTCTTGGAAAAAACAATAGCTACTCCCTCCGCTCCCTCTGAAATGT	1136
	Q	y	1141	ATACATATGGATTGGACACGAGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG	1200
	D	b	1137	ATACATATGGATTGGACACGAGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG	1196
	Q	y	1201	AAAGAGAAAGAAAGTCGGTTAAAGTACGGGACCACCAATATATATAATTCATAGATTTAG	1260
	D	b	1197	AAAGAGAAAGAAAGTCGGTTAAAGTACGGGACCACCAATATATATAATTCATAGATTTAG	1256
	Q	y	1261	AAAAAGTAGTTGAAAAGTAGTGGGTGGGATTTTTATATATATAAAAAATTTACTATTTTTG	1320
	D	b	1257	AAAAAGTAGTTGAAAAGTAGTGGGTGGGATTTTTATATATATAAAAAATTTACTATTTTTG	1316
	Q	y	1321	AGAAAGTTTTGAAATGTATAGAAATGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT	1380
	D	b	1317	AGAAAGTTTTGAAATGTATAGAAATGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAT	1376
	Q	y	1381	TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTTGTATTCCA	1440
	D	b	1377	TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTTGTATTCCA	1436
	Q	y	1441	TAGATTATAAATCTATGTTTATAATGATATATAAATTTTAAAAATACTATATAATTAATT	1500
	D	b	1437	TAGATTATAAATCTATGTTTATAATGATATATAAATTTTAAAAATACTATATAATTAATT	1496
	Q	y	1501	CTGATTAGTCGATTACCGCCCTTTTATAAATTTTACAATCTAGTAAATATGAATAAAATCAG	1560
	D	b	1497	CTGATTAGTCGATTACCGCCCTTTTATAAATTTTACAATCTAGTAAATATGAATAAATCAG	1556
	Q	y	1561	TTATCTGAAAGCAAATAATATCTTTGTFAAACAGCGTTCGGTCAAAATGGGAAGTTCCATG	1620
	D	b	1557	TTATCTGAAAGCAAATAATATCTTTGTFAAACAGCGTTCGGTCAAAATGGGAAGTTCCATG	1616
	Q	y	1621	TGTTATCAATAGTTTTTATAATAAAGTAATAATTTTAAATTAATTTTGTGTTATTTTGTTCAGA	1680
	D	b	1617	TGTTATCAATAGTTTTTATAATAAAGTAATAATTTTAAATTAATTTTGTGTTATTTTGTTCAGA	1676
	Q	y	1681	AATTTAAAAATAAATTTATTTGAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACT	1740
	D	b	1677	AATTTAAAAATAAATTTATTTGAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACT	1736
	Q	y	1741	GTTTGAACAATGTATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATPAAT	1800

Db	1737	GTTTGAAACAATGATGTGTCGGGTGTCACATCTATGACCTTTTCAACTCTAAACTAGTAGGAATAAT	1796
Qy	1801	GCATCTCTAGAAATACATCTTTTCAAATTTTCAACAAAACACAGCTTTTAACTTTTCTTTTCAACG	1860
Db	1797	GCATCTCTAGAAATACATCTTTTCAAATTTTCAACAAAACACAGCTTTTAACTTTTCTTTTCAACG	1856
Qy	1861	GATTGGNAATCCTTTTCTAAACCTTTTAAAAATAAAAAATGCAATTAATGTGAATATTATTC	1920
Db	1857	GATTGGNAATCCTTTTCTAAACCTTTTAAAAATAAAAAATGCAATTAATGTGAATATTATTC	1916
Qy	1921	AACACCTTCAACATTTGATGTTTACGGTACTATAAATAGTGCTCTTGGTGCTCTACTATTCAT	1980
Db	1917	AACACCTTCAACATTTGATGTTTACGGTACTATAAATAGTGCTCTTGGTGCTCTACTATTCAT	1976
Qy	1981	CACATCAATCTTTACACACAAAACCTTGAGCGTTAAATTTTTTCTACTATTTCTCAGCAAT	2037
Db	1977	CACATCAATCTTTACACACAAAACCTTGAGCGTTAAATTTTTTCTACTATTTCTCAGCAAT	2033

RESULT 6
US-09-300-487-1
; Sequence 1, Application US/09300487
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,434
; FILING DATE: 12-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..247
; US-09-300-487-1

Query Match 2.8%; Score 57; DB 20; Length 247;
Best Local Similarity 100.0%; Pred. No. 6.8e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2.8%; Score 57; DB 20; Length 247;

Best Local Similarity 100.0%; Pred. No. 6.8e-16;

Matches	57	Conservative	0	Mismatches	0	Indels	0	Gaps	0
---------	----	--------------	---	------------	---	--------	---	------	---

Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCATCAATCTTT 1992
Db 131 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCATCAATCTTT 187

RESULT 7

US-09-300-487-2
; Sequence 2, Application US/09300487
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,434
; FILING DATE: 12-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..2042
US-09-300-487-2

Query Match 2.8%; Score 57; DB 20; Length 2042;
Best Local Similarity 100.0%; Pred.No. 7.7e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1936 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCATCAATCTTT 1992
Db 1926 ATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTACTATCATCATCAATCTTT 1982

Search completed: October 28, 2005, 07:48:10
Job time : 8051 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 03:19:39 ; Search time 6540 Seconds

(without alignments)
4511.243 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtccctacagcaca.....gcaataacattctaataatc 2052

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 27723106 seqs, 7189968421 residues

Word size : 50

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2.*
- 3: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1.*
- 8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2.*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq10.*
- 11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq11.*
- 12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq12.*
- 13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq13.*
- 14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq14.*
- 15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq15.*
- 16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq16.*
- 17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq17.*
- 18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq18.*
- 19: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq19.*
- 20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2.*
- 22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3.*
- 23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq4.*
- 24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq5.*
- 25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq6.*
- 26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq7.*
- 27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq8.*
- 28: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	841	41.0	2930	27	US-11-130-391-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-11-130-391-5
; Sequence 5, Application US/11130391
; GENERAL INFORMATION:
; APPLICANT: Sumitomo Chemical Co., Ltd.
; TITLE OF INVENTION: HERBICIDAL COMPOUND RESISTANT PLANT
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/130.391
; PRIOR FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: JP 2004/147363
; PRIOR FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: JP 2005/070980
; PRIOR FILING DATE: 2005-03-14
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930)...(2561)
; FEATURE:
; OTHER INFORMATION: Part of an expression plasmid containing a Protoporphyrinogen IX
US-11-130-391-5

Query Match 41.0%; Score 841; DB 27; Length 2930;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1146	TATGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAAGA	1205
Db	15	TATGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAAGA	74
Qy	1206	GAAGAAAAAGTGGTAAAGTAGCCGGACCCACCAATATATATTAATGTAGATTTAGAAAAG	1265
Db	75	GAAGAAAAAGTGGTAAAGTAGCCGGACCCACCAATATATATTAATGTAGATTTAGAAAAG	134
Qy	1266	TAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTGAGAAA	1325
Db	135	TAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTGAGAAA	194
Qy	1326	GTTTTGAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAAGTGTATAGAAATTAAT	1385
Db	195	GTTTTGAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAAGTGTATAGAAATTAAT	254
Qy	1386	GGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGTATTTTCATAAGA	1445
Db	255	GGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGTATTTTCATAAGA	314
Qy	1446	TTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAAATTCAT	1505
Db	315	TTATAAATCTATGTTATAATGATAATATAATTTTAAAAATAATACTATATTAAATTCAT	374
Qy	1506	TAGTCGATTACCGCTTTTATAAATTTTACAATCTGAGTAAATGAATAAATCAAGTTATC	1565
Db	375	TAGTCGATTACCGCTTTTATAAATTTTACAATCTGAGTAAATGAATAAATCAAGTTATC	434
Qy	1566	TGAAGCAATAATATCTTTGTAAGAACAGCGTTCGCTCAATCGGAAGTTCACTGTGTAT	1625
Db	435	TGAAGCAATAATATCTTTGTAAGAACAGCGTTCGCTCAATCGGAAGTTCACTGTGTAT	494
Qy	1626	TCAATAGTTTTTAATAAAGTAAATTTTAAATTAATTTTGTATTTTGTATTTTGTATTTT	1685
Db	495	TCAATAGTTTTTAATAAAGTAAATTTTAAATTAATTTTGTATTTTGTATTTTGTATTTT	554
Qy	1686	AAAAATAAATTTAGAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACTGTTG	1745
Db	555	AAAAATAAATTTAGAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACTGTTG	614
Qy	1746	AACAATGTATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAATGCATT	1805
Db	615	AACAATGTATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAATGCATT	674

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 07:54:14 ; Search time 1471 Seconds
(without alignments)
988.210 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 acaatgtatgtccggtgacatcatgatgac 30

Scoring table: IDENTITY_NUC

Gapop_10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hig.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pt.*

10: gb_ro.*

11: gb_sts.*

12: gb_by.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	2048	6	E40089
2	30	100.0	2048	6	E40090
3	30	100.0	2052	6	E40087
4	30	100.0	2052	6	E40093
5	30	100.0	2056	6	E40091
6	23.6	78.7	2042	6	AR076817
7	23.6	78.7	2042	6	AR076817
8	21.8	72.7	348283	2	CR352244
9	21.2	70.7	161762	2	CR352244
10	21.2	70.7	162325	2	CR774195
11	21.2	70.7	213033	2	AC110515
12	21	70.0	140142	8	AC146937
13	21	70.0	185386	10	AC025586
14	21	70.0	205668	10	AL592422
15	20.6	68.7	69674	9	AL136130
16	20.6	68.7	78026	2	AC023516
17	20.6	68.7	215896	2	AC046180
18	20.4	68.0	118001	9	AC073081
19	20.4	68.0	193910	2	AC040996

C	20	68.0	194235	9	AC109822	Homo sapi
C	21	68.0	288385	2	AC092430	Homo sapi
	22	67.3	121101	2	AC008034	Homo sapi
	23	67.3	155313	2	AC068315	Homo sapi
	24	67.3	157007	2	AC066599	Homo sapi
C	25	67.3	161903	2	AC021996	Homo sapi
C	26	67.3	185067	9	AC022382	Homo sapi
C	27	67.3	189430	2	AC011610	Homo sapi
C	28	67.3	197360	2	AC018829	Homo sapi
C	29	67.3	216530	2	EX649535	Danio rer
	30	67.3	260792	2	AC125770	Rattus no
C	31	67.3	320398	2	AC136192	Rattus no
	32	66.7	81398	2	AC023821	Homo sapi
	33	66.7	103152	2	AP004091	Oryza sat
	34	66.7	141322	2	AC139637	Mus muscu
	35	66.7	158601	8	OSJN00059	
	36	66.7	163328	8	AP004865	Oryza sat
C	37	66.7	172879	10	AC142474	Mus muscu
C	38	66.7	186669	2	AC151188	Bos tauru
C	39	66.7	201294	2	AC128748	Rattus no
	40	66.7	204302	2	AC068650	Mus muscu
	41	66.7	216935	2	AC132648	Rattus no
C	42	66.7	234787	2	AC073780	Mus muscu
C	43	66.7	244520	2	AC128555	Rattus no
	44	66.7	245754	2	AC127884	Rattus no
	45	66.7	254067	2	AC133699	Rattus no
	46	66.7	256396	2	AC108524	Rattus no
C	47	66.7	315370	2	AC109390	Rattus no
C	48	65.3	2931	6	Q0804634	Sequence
C	49	65.3	99231	8	AC027665	Genomic s
	50	65.3	110000	2	AC095866	Rattus no
	51	65.3	111153	10	AL928993	Mouse DNA
	52	65.3	143879	8	AC069251	Genomic s
C	53	65.3	158335	9	AL365199	Human DNA
	54	65.3	174620	2	AC119233	Mus muscu
	55	65.3	178036	2	AP001188	Homo sapi
C	56	65.3	178059	9	AC092706	Homo sapi
C	57	65.3	181054	2	AC130189	Pan trogl
C	58	65.3	183155	2	AL591169	Homo sapi
C	59	65.3	203942	2	AC117760	Mus muscu
C	60	65.3	204733	2	AC139299	Mus muscu
	61	65.3	216824	2	AC110116	Rattus no
	62	65.3	231363	2	AC107804	Mus muscu
	63	65.3	231842	2	EX510341	Mus muscu
C	64	65.3	231842	2	EX510341	Mus muscu
C	65	65.3	262038	5	EX511227	Zebrafish
	66	64.7	6081	9	HSM806528	Strongylo
C	67	64.7	61857	2	AC149981	Strongylo
	68	64.7	95345	9	AC007029	Homo sapi
	69	64.7	134319	9	AC113405	Homo sapi
	70	64.7	136819	10	AC138382	Mus muscu
C	71	64.7	141122	2	AC149978	Strongylo
	72	64.7	148847	2	AC023101	Homo sapi
C	73	64.7	155305	2	AC117902	Rattus no
C	74	64.7	160433	9	AC093296	Homo sapi
C	75	64.7	160703	9	AC022104	Homo sapi
	76	64.7	175928	2	AC144859	Mus muscu
	77	64.7	178046	5	AL772158	Zebrafish
	78	64.7	178161	2	AC150825	Callithri
	79	64.7	179424	2	AC027222	Homo sapi
	80	64.7	181202	2	AC103623	Mus muscu
C	81	64.7	187925	2	AC103361	Mus muscu
	82	64.7	192814	2	AC114585	Mus muscu
	83	64.7	193766	2	AC090290	Mus muscu
C	84	64.7	197694	2	AC136065	Rattus no
	85	64.7	201783	10	AC114410	Mus muscu
	86	64.7	202234	2	AC118815	Rattus no
	87	64.7	202856	2	AC027374	Homo sapi
C	88	64.7	203580	2	EX649256	Danio rer
	89	64.7	208035	2	AC134450	Mus muscu
C	90	64.7	216687	2	AC119913	Mus muscu
C	91	64.7	218337	10	AC124532	Mus muscu
	92	64.7	237515	2	AC126059	Rattus no

93 19.4 64.7 238283 2 AC094520
 c 94 19.4 64.7 240932 2 AC098206
 c 95 19.4 64.7 250272 2 AC111902
 96 19.4 64.7 251930 10 AL772293
 97 19.4 64.7 272085 2 AC098400
 98 19.2 64.0 2384 5 CR760547
 c 99 19.2 64.0 13352 1 AE003947
 100 19.2 64.0 67211 2 AC133284

ALIGNMENTS

RESULT 1
 E40089
 LOCUS 2048 bp DNA linear PAT 31-JAN-2002
 DEFINITION Plant promoter and terminator.
 ACCESSION E40089
 VERSION E40089.1 GI:18627205
 KEYWORDS JP 2000166577-A/3.
 SOURCE unidentified
 ORGANISM unidentified

REFERENCE 1 (bases 1 to 2048)
 AUTHORS Nishikawa,S. and Oeda,K.
 TITLE Plant promoter and terminator
 JOURNAL Patent: JP 2000166577-A 3 20-JUN-2000;
 SUMITOMO CHEM CO LTD
 COMMENT OS Daucus carota L.
 PN JP 2000166577-A/3
 PD 20-JUN-2000
 PF 01-OCT-1999 JP 1999281475

PI SATOMI NISHIKAWA, KENJI OEDA
 PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1:91), PC
 C12N15/00,
 PC C12N5/00, (C12N5/00,C12R1:91)
 CC
 FH Key Location/Qualifiers
 FT promoter (1)..(2048).
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source
 1..2048
 Location/Qualifiers
 FT promoter (1)..(2048).
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 30; DB 6; Length 2048;
 Best Local Similarity 100.0%; Pred. No. 0.00049;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGATCATCTATGAC 30
 |||||
 Db 1742 AACAAATGATGTCGGGTGATCATCTATGAC 1771

RESULT 2
 E40090
 LOCUS 2048 bp DNA linear PAT 31-JAN-2002
 DEFINITION Plant promoter and terminator.
 ACCESSION E40090
 VERSION E40090.1 GI:18627206
 KEYWORDS JP 2000166577-A/4.
 SOURCE unidentified
 ORGANISM unidentified

REFERENCE 1 (bases 1 to 2048)
 AUTHORS Nishikawa,S. and Oeda,K.
 TITLE Plant promoter and terminator
 JOURNAL Patent: JP 2000166577-A 4 20-JUN-2000;
 SUMITOMO CHEM CO LTD
 COMMENT OS Daucus carota L.
 PN JP 2000166577-A/4

PD 20-JUN-2000
 PF 01-OCT-1999 JP 1999281475
 PR SATOMI NISHIKAWA, KENJI OEDA
 PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1:91), PC
 C12N15/00,
 PC C12N5/00, (C12N5/00,C12R1:91)
 CC
 FH Key Location/Qualifiers
 FT promoter (1)..(2048).
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source
 1..2048
 Location/Qualifiers
 FT promoter (1)..(2048).
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 30; DB 6; Length 2048;
 Best Local Similarity 100.0%; Pred. No. 0.00049;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGATCATCTATGAC 30
 |||||
 Db 1742 AACAAATGATGTCGGGTGATCATCTATGAC 1771

RESULT 3
 E40087
 LOCUS 2052 bp DNA linear PAT 31-JAN-2002
 DEFINITION Plant promoter and terminator.
 ACCESSION E40087
 VERSION E40087.1 GI:18627203
 KEYWORDS JP 2000166577-A/1.
 SOURCE unidentified
 ORGANISM unidentified

REFERENCE 1 (bases 1 to 2052)
 AUTHORS Nishikawa,S. and Oeda,K.
 TITLE Plant promoter and terminator
 JOURNAL Patent: JP 2000166577-A 1 20-JUN-2000;
 SUMITOMO CHEM CO LTD
 COMMENT OS Daucus carota L.
 PN JP 2000166577-A/1
 PD 20-JUN-2000
 PF 01-OCT-1999 JP 1999281475

PI SATOMI NISHIKAWA, KENJI OEDA
 PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1:91), PC
 C12N15/00,
 PC C12N5/00, (C12N5/00,C12R1:91)
 CC
 FH Key Location/Qualifiers
 FT promoter (1)..(2052).
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source
 1..2052
 Location/Qualifiers
 FT promoter (1)..(2052).
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 30; DB 6; Length 2052;
 Best Local Similarity 100.0%; Pred. No. 0.00049;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGATCATCTATGAC 30
 |||||
 Db 1746 AACAAATGATGTCGGGTGATCATCTATGAC 1775

RESULT 4
 E40093
 LOCUS 2052 bp DNA linear PAT 31-JAN-2002
 DEFINITION Plant promoter and terminator.

ACCESSION E40093
VERSION E40093.1 GI:18627209
KEYWORDS JP 2000166577-A/7.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2052)
AUTHORS Nishikawa, S. and Oeda, K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 7 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000166577-A/7
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC
C12N15/00,
PC C12N5/00, (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(2052).
FEATURES source
1..2052
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 30; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACATGTATGTCGGGTACATCTATGAC 30
|||||
Db 1746 AACAAATGTATGTCGGGTACATCTATGAC 1775
|||||
RESULT 5
E40091
LOCUS E40091 2056 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40091
VERSION E40091.1 GI:18627207
KEYWORDS JP 2000166577-A/5.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2056)
AUTHORS Nishikawa, S. and Oeda, K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 5 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000166577-A/5
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC
C12N15/00,
PC C12N5/00, (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(2056).
FEATURES source
1..2056
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACATGTATGTCGGGTACATCTATGAC 30
|||||
Db 1746 AACAAATGTATGTCGGGTACATCTATGAC 1775
|||||
RESULT 6
AR076817
LOCUS AR076817 2042 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959176.
ACCESSION AR076817
VERSION AR076817.1 GI:10003563
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai, S. and Oeda, K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 2 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..2042
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 78.7%; Score 23.6; DB 6; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.95;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AACATGTATGTCGGGTACATCTATGAC 30
|||||
Db 1737 AACAACTTTGTCGGGTATATTATGAC 1766
|||||
RESULT 7
E15125
LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai, T. and Oita, K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (Carrot)
PN JP 1998052273-A/2
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09, A01H5/00, C07H21/04, C07K14/415, C12N1/21, C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..2042
/organism="Daucus carota L."
FT promoter 1<..<2042.
FEATURES Location/Qualifiers
source 1..2042
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"


```

/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-11J9"
1. .1925
/notes="wgs end_extension
clone_end:sp6"
1909. .2527
/notes="clone_boundary
clone_end:sp6
site:
end_sequence:BH342645"
9272. .11765
/notes="wgs_contig"
11866. .15419
/notes="wgs_contig"
20601. .22623
/notes="wgs_contig"
22674. .27148
/notes="wgs_contig"
48968. .52962
/notes="wgs_contig"
53013. .54143
/notes="wgs_contig"
54544. .56816
/notes="wgs_contig"
56870. .59602
/notes="wgs_contig"
59703. .61990
/notes="wgs_contig"
68117. .71973
/notes="wgs_contig"
338673. .340201
/notes="wgs_contig"
340622. .343138
/notes="wgs_contig"

```

ORIGIN

```

Query Match      72.7%; Score 21.8; DB 2; Length 348283;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 AACATGATGTCGGGTACATCT 25
|||||

```

```

Db 90808 AACATGATGTCAGGTGACGTCT 90832
|||||

```

RESULT 9

```

CR352244/c      161762 bp      DNA      linear      HTG 27-MAR-2004
LOCUS          Danio rerio clone CH211-131E11, *** SEQUENCING IN PROGRESS ***, 14
DEFINITION      unordered pieces.

```

```

ACCESSION      CR352244
VERSION         CR352244.4 GI:45772227
KEYWORDS        HTG; HTGS PHASE1.
SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 161762)
Sims,S.

```

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

Submitted (26-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 27, 2004 this sequence version replaced gi:45598726.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: zC131E11

```

----- Summary Statistics

```

Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 157149 bases at least Q40
Consensus quality: 157757 bases at least Q30
Consensus quality: 158375 bases at least Q20
Insert size: 160462; sum-of-contigs
Insert size: 174990; 4.8% error; agarose-fp
Quality coverage: 8.14x in Q20 bases; sum-of-contigs Quality
coverage: 7.67x in Q20 bases; agarose-fp

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 16722: contig of 16722 bp in length
* 16723 16822: gap of 100 bp
* 16823 28393: contig of 11571 bp in length
* 28394 40368: contig of 11875 bp in length
* 40369 40468: gap of 100 bp
* 40469 57841: contig of 17373 bp in length
* 57842 57941: gap of 100 bp
* 57942 69071: contig of 11130 bp in length
* 69072 69171: gap of 100 bp
* 69172 77218: contig of 8047 bp in length
* 77219 85619: contig of 8301 bp in length
* 85620 85719: gap of 100 bp
* 85720 89373: contig of 3654 bp in length
* 89374 89473: gap of 100 bp
* 89474 98585: contig of 9112 bp in length
* 98586 103814: gap of 100 bp
* 103815 103914: contig of 5129 bp in length
* 103915 110517: contig of 6603 bp in length
* 110518 110617: gap of 100 bp
* 110618 139606: contig of 28989 bp in length
* 139607 139706: gap of 100 bp
* 139707 144556: contig of 4850 bp in length
* 144557 144657: gap of 100 bp
* 144658 161762: contig of 17106 bp in length.

```

FEATURES

source

```

1. .161762
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-131E11"
/clone_lib="CHORI-211"
1. .16722
/notes="assembly fragment:01696
fragment_chain:1
clone_end:SP6
vector side:left"
16823. .28393
/notes="assembly fragment:00818
fragment_chain:1"
28494. .40368
/notes="assembly fragment:00996
fragment_chain:1"
40469. .57841
/notes="assembly fragment:01195
fragment_chain:1"
57942. .59071
/notes="assembly fragment:00668
fragment_chain:1"
69172. .77218
/notes="assembly fragment:00429
fragment_chain:1"

```

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature


```

misc_feature 77319..85619
/notes="assembly_fragment:00216
fragment_chain:1"
misc_feature 85720..89373
/notes="assembly_fragment:00050
fragment_chain:1"
misc_feature 89474..98585
/notes="assembly_fragment:00542
fragment_chain:1"
misc_feature 98686..103814
/notes="assembly_fragment:00090
fragment_chain:1"
misc_feature 103915..110517
/notes="assembly_fragment:00319
fragment_chain:1"
misc_feature 110618..1139606
/notes="assembly_fragment:02011
fragment_chain:2"
misc_feature 139707..144556
/notes="assembly_fragment:00147
fragment_chain:2"
misc_feature 144657..161762
/notes="assembly_fragment:01414
fragment_chain:2
clone_end:T7
vector_side:right"

ORIGIN
Query Match 70.7%; Score 21.2; DB 2; Length 161762;
Best Local Similarity 88.5%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 AATGTATGTCGGGTGCATCTATGA 29
|||||
Db 9417 AATGTGTCCGGGTGCATCTGTGA 9392

RESULT 10
CR774195 162325 bp DNA linear HTG 16-SEP-2004
LOCUS
DEFINITION Danio rerio clone DKEY-73P2, *** SEQUENCING IN PROGRESS ***, 7
unordered pieces.
ACCESSION CR774195
VERSION CR774195.1 GI:52213980
KEYWORDS HTG; HTGS PHASE1.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 162325)
Mclay, K.
Direct Submission
Submitted (15-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK73P2
----- Summary Statistics
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 160506 bases at least Q40
Consensus quality: 160914 bases at least Q30
Consensus quality: 161069 bases at least Q20
Insert size: 161725; sum-of-contigs
Insert size: 167136; 0.4% error; agarose-fp
Quality coverage: 8.68x in Q20 bases; sum-of-contigs Quality
coverage: 8.41x in Q20 bases; agarose-fp

```

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 17481: contig of 17481 bp in length
* 17482 17581: gap of 100 bp
* 17582 38294: contig of 20713 bp in length
* 38295 38394: gap of 100 bp
* 38395 116218: contig of 77824 bp in length
* 116219 116318: gap of 100 bp
* 116319 119531: contig of 3213 bp in length
* 119532 119631: gap of 100 bp
* 119632 155280: contig of 35649 bp in length
* 155281 155380: gap of 100 bp
* 155381 159571: contig of 4191 bp in length
* 159572 159671: gap of 100 bp
* 159672 162325: contig of 2654 bp in length.

```

FEATURES

```

source
1..162325
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-73P2"
/clone_lib="DanioKey"
1..17481
/notes="assembly_fragment:00118
fragment_chain:1"
17582..38294
/notes="assembly_fragment:00360
fragment_chain:1"
38395..116218
/notes="assembly_fragment:01043
fragment_chain:1"
116319..119531
/notes="assembly_fragment:00044
fragment_chain:1"
119632..155280
/notes="assembly_fragment:00616
fragment_chain:1"
155381..159571
/notes="assembly_fragment:00075
fragment_chain:1"
159672..162325
/notes="assembly_fragment:00014.0"

```

ORIGIN

```

Query Match 70.7%; Score 21.2; DB 2; Length 162325;
Best Local Similarity 88.5%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 AATGTATGTCGGGTGCATCTATGA 29
|||||
Db 100707 AATGTGTCCGGGTGCATCTGTGA 100732

```

RESULT 11

```

AC110515 213033 bp DNA linear HTG 09-MAR-2004
LOCUS
DEFINITION Mus musculus chromosome 13 clone RP23-295C1 map 13, WORKING DRAFT
SEQUENCE, 8 ordered pieces.
ACCESSION AC110515
VERSION AC110515.4 GI:44986734
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 213033)

```


ACCESSION ACL46937
 VERSION ACL46937.2 GI:48958697
 KEYWORDS HTG.
 SOURCE
 ORGANISM *Oryza sativa* (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoae; *Oryza*.
 REFERENCE 1 (bases 1 to 140142)
 AUTHORS Buell, C.R., Yuan, Q., Cuiyang, S., Liu, J., Gansberger, K., Jones, K.M., Overton II, L., Tsifrin, T., Kim, M., Bera, J., Jin, S., Fadrosch, D.W., Tallon, L., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S., Riedmuller, S.B., Uterbach, T., Feldblum, T., Yang, Q., Haas, B., Suh, B., Peterson, J., Quackenbush, J., White, O., Salzberg, S. and Fraser, C.
 TITLE *Oryza sativa* sep. japonica cv. Nipponbare OSJNba004015 BAC genomic sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 140142)
 AUTHORS Buell, R.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 REFERENCE 3 (bases 1 to 140142)
 AUTHORS Buell, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 REFERENCE 4 (bases 1 to 140142)
 AUTHORS Buell, R.
 TITLE Direct Submission
 JOURNAL Submitted (19-JUN-2004) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
 COMMENT On Jun 19, 2004 this sequence version replaced gi:37991935.
 FEATURES
 source
 1..140142
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="11"
 /clone="OSJNba004015"
 /note="japonica cultivar-group"
 ORIGIN
 Query Match 70.0%; Score 21; DB 8; Length 140142;
 Best Local Similarity 82.8%; Pred. No. 29;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 2 ACAATGATGTCGGGTGATCATGAC 30
 Db 113334 ACAATATTTGTTGGTGTACATGATGAC 113306
 RESULT 13
 AC025586
 LOCUS AC025586 185386 bp DNA linear ROD 24-JUL-2002
 DEFINITION Genomic sequence for Mus musculus, clone RP23-320C8, complete sequence.
 VERSION AC025586
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 185386)
 AUTHORS McCombie, W.R., de la Bastide, M., Spiegel, L., Preston, R., Nascimben, L., Zutavern, T., Ballja, V., Bell, M., Miller, B., Katzenberger, F., Muller, S., Sullivan, P., Yang, C., Dike, S., Palmer, L., O'Shaughnessy, A. and Dedhia, N.
 TITLE Genomic sequence for Mus musculus, clone RP23-320C8, complete sequence

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 185386)
 AUTHORS McCombie, W.R.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAR-2000) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
 REFERENCE 3 (bases 1 to 185386)
 AUTHORS McCombie, W.R.
 TITLE Direct Submission
 JOURNAL Submitted (24-JUL-2002) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
 COMMENT On Jul 24, 2002 this sequence version replaced gi:9972313.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.
 FEATURES
 Location/Qualifiers
 1..185386
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-320C8"
 /clone_lib="RPCI-23"
 complement(23305..23539)
 /note="We believe the assembly to be correct. The sequence is a simple repeat (TTCCC)n in which the exact number of repeat copies is unknown. One high quality subclone in the region spans the repeat into unique sequence on both sides."
 misc_feature
 Query Match 70.0%; Score 21; DB 10; Length 185386;
 Best Local Similarity 82.8%; Pred. No. 30;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 AACATGTATGTCGGGTGATCATGATGA 29
 Db 98187 AACATGTATGTCGGTTAGATCTAAGA 98215
 RESULT 14
 AL592422
 LOCUS AL592422 205668 bp DNA linear ROD 16-FEB-2002
 DEFINITION Mouse DNA sequence from clone RP23-218016 on chromosome 11, complete sequence.
 VERSION AL592422
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 205668)
 AUTHORS Almeida, J.
 TITLE Direct Submission
 JOURNAL Submitted (16-FEB-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT On Feb 21, 2002 this sequence version replaced gi:17426517.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivari, T.M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6978142.
All repeats were identified using RepeatMasker:

COMMENT

Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1909

Center clone name: 115_O_4

* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
800: contig of 800 bp in length
* 801: gap of 100 bp
* 1749: contig of 849 bp in length
* 1750: gap of 100 bp
* 1850: contig of 824 bp in length
* 2674: gap of 100 bp
* 2774: contig of 841 bp in length
* 3614: gap of 100 bp
* 3714: contig of 808 bp in length
* 3715: gap of 100 bp
* 4522: contig of 833 bp in length
* 4623: gap of 100 bp
* 4623: contig of 833 bp in length
* 5455: gap of 100 bp
* 5456: contig of 833 bp in length
* 5556: gap of 100 bp
* 6388: contig of 819 bp in length
* 6389: gap of 100 bp
* 6489: contig of 819 bp in length
* 7307: gap of 100 bp
* 7407: contig of 819 bp in length
* 7308: gap of 100 bp
* 8226: contig of 819 bp in length
* 8327: gap of 100 bp
* 9160: contig of 834 bp in length
* 9260: gap of 100 bp
* 9261: contig of 804 bp in length
* 10064: contig of 804 bp in length
* 10164: gap of 100 bp
* 10065: contig of 832 bp in length
* 10165: gap of 100 bp
* 10996: contig of 814 bp in length
* 10997: gap of 100 bp
* 11097: contig of 814 bp in length
* 11911: gap of 100 bp
* 12011: contig of 845 bp in length
* 12855: gap of 100 bp
* 12856: contig of 845 bp in length
* 12956: gap of 100 bp
* 13801: contig of 830 bp in length
* 13901: gap of 100 bp
* 14731: contig of 845 bp in length
* 14831: gap of 100 bp
* 15676: contig of 835 bp in length
* 16610: gap of 100 bp
* 16711: contig of 829 bp in length
* 17539: gap of 100 bp
* 17540: contig of 819 bp in length
* 17474: contig of 835 bp in length
* 18574: gap of 100 bp
* 18575: contig of 815 bp in length
* 19389: gap of 100 bp
* 19489: contig of 839 bp in length
* 20328: gap of 100 bp
* 20428: contig of 848 bp in length
* 21276: gap of 100 bp
* 21377: contig of 798 bp in length
* 22174: gap of 100 bp
* 22275: contig of 839 bp in length
* 23113: gap of 100 bp
* 23214: contig of 844 bp in length
* 24058: gap of 100 bp
* 24158: contig of 849 bp in length
* 25007: gap of 100 bp
* 25107: contig of 829 bp in length
* 25936: gap of 100 bp
* 26036: contig of 835 bp in length
* 26871: gap of 100 bp
* 26970: contig of 842 bp in length
* 27812: gap of 100 bp
* 27913: contig of 823 bp in length
* 28735: gap of 100 bp
* 28836: contig of 838 bp in length
* 29674: gap of 100 bp
* 29774: contig of 823 bp in length
* 30597: gap of 100 bp
* 30597: contig of 845 bp in length
* 31541: gap of 100 bp
* 31542: contig of 848 bp in length
* 32490: gap of 100 bp
* 32590: contig of 789 bp in length
* 33379: gap of 100 bp
* 34333: contig of 854 bp in length
* 34333: gap of 100 bp
* 35294: contig of 862 bp in length
* 35394: gap of 100 bp
* 35395: contig of 845 bp in length
* 36239: gap of 100 bp
* 36340: contig of 822 bp in length
* 37161: gap of 100 bp
* 37261: contig of 826 bp in length
* 38087: gap of 100 bp
* 38188: contig of 823 bp in length
* 39011: gap of 100 bp
* 39111: contig of 837 bp in length
* 39947: gap of 100 bp
* 40048: contig of 839 bp in length
* 40887: gap of 100 bp
* 40887: contig of 843 bp in length
* 41830: gap of 100 bp
* 41830: contig of 820 bp in length
* 42750: gap of 100 bp
* 42750: contig of 838 bp in length
* 43688: gap of 100 bp
* 43688: contig of 833 bp in length
* 44621: gap of 100 bp
* 44621: contig of 848 bp in length
* 45569: gap of 100 bp
* 45569: contig of 827 bp in length
* 46496: gap of 100 bp
* 46496: contig of 826 bp in length
* 47422: gap of 100 bp
* 47422: contig of 824 bp in length
* 48346: gap of 100 bp
* 48346: contig of 841 bp in length
* 49287: gap of 100 bp
* 49287: contig of 812 bp in length
* 50199: gap of 100 bp
* 50199: contig of 847 bp in length
* 51145: gap of 100 bp
* 51145: contig of 819 bp in length
* 51246: contig of 819 bp in length

```

* 52065 52164: gap of 100 bp
* 52165 contig of 842 bp in length
* 53006 contig of 100 bp
* 53107 contig of 823 bp in length
* 53929 contig of 100 bp
* 53930 contig of 806 bp in length
* 54030 contig of 100 bp
* 54936 contig of 824 bp in length
* 54936 contig of 100 bp
* 55759 contig of 822 bp in length
* 55859 contig of 100 bp
* 56681 contig of 826 bp in length
* 56781 contig of 100 bp
* 56782 contig of 826 bp in length
* 57607 contig of 100 bp
* 57707 contig of 830 bp in length
* 58537 contig of 100 bp
* 58637 contig of 819 bp in length
* 59456 contig of 100 bp
* 59556 contig of 845 bp in length
* 60401 contig of 100 bp
* 60501 contig of 844 bp in length
* 61345 contig of 100 bp
* 61346 contig of 819 bp in length
* 62264 contig of 100 bp
* 62364 contig of 830 bp in length
* 63194 contig of 100 bp
* 63294 contig of 834 bp in length
* 64128 contig of 100 bp
* 64228 contig of 100 bp

Query Match      68.7%; Score 20.6; DB 2; Length 78026;
Best Local Similarity 85.2%; Pred. No. 44;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACAATGATGTCGGTGTCATCATG 28
Db 40150 AAACGATGTCAGTGACATTTATG 40176

RESULT 17
AC046180
LOCUS      215896 bp DNA linear HTG 12-MAY-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-465J17 map 15, WORKING DRAFT
SEQUENCE   27 unordered pieces.
AC046180
VERSION    AC046180.2 GI:7770547
KEYWORDS   HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,P.,
            Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,W., Collins,S.,
            Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
            Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marguis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Poyan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,E., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

```

```

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 12, 2000 this sequence version replaced gi:7547220.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8718
Center clone name: 465_J_17
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 203149 bases at least Q40
Consensus quality: 209316 bases at least Q30
Consensus quality: 211902 bases at least Q20
Insert size: 220000; agarose-fp
Insert size: 213296; sum-of-contigs
Quality coverage: 4.3 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1011: contig of 1011 bp in length
* 1012 1111: gap of 100 bp
* 1112 2633: contig of 1522 bp in length
* 2634 2734: gap of 100 bp
* 2734 4093: contig of 1360 bp in length
* 4094 4193: gap of 100 bp
* 4194 5298: contig of 1105 bp in length
* 5299 5398: gap of 100 bp
* 5399 7088: contig of 1690 bp in length
* 7089 7189: gap of 100 bp
* 7189 7890: contig of 701 bp in length
* 7890 10304: contig of 2315 bp in length
* 10305 10405: gap of 100 bp
* 10405 12955: contig of 2551 bp in length
* 12956 13056: gap of 100 bp
* 13056 16427: contig of 3371 bp in length
* 16427 19178: gap of 100 bp
* 19179 19278: contig of 2652 bp in length
* 19279 23519: gap of 100 bp
* 23520 23619: contig of 4241 bp in length
* 23619 26853: gap of 100 bp
* 26854 26953: contig of 3234 bp in length
* 26954 31292: gap of 100 bp
* 31293 31392: contig of 4339 bp in length
* 31393 36401: contig of 5009 bp in length
* 36402 36501: gap of 100 bp
* 36502 41321: contig of 4820 bp in length
* 41322 41421: gap of 100 bp
* 41422 47286: contig of 5865 bp in length
* 47287 47386: gap of 100 bp
* 47387 54497: contig of 7111 bp in length
* 54498 54598: gap of 100 bp
* 54598 60757: contig of 6160 bp in length
* 60758 60857: gap of 100 bp

```

TITLE
JOURNAL
COMMENT

```

* 60858 68004: contig of 7147 bp in length
* 68005 68104: gap of 100 bp
* 68105 77167: contig of 9063 bp in length
* 77168 77267: gap of 100 bp
* 77268 88567: contig of 11300 bp in length
* 88568 88667: gap of 100 bp
* 88668 101419: contig of 12752 bp in length
* 101420 101519: gap of 100 bp
* 101520 114909: contig of 13390 bp in length
* 114910 115009: gap of 100 bp
* 115010 130182: contig of 15173 bp in length
* 130183 130282: gap of 100 bp
* 130283 151503: contig of 21221 bp in length
* 151504 151604: gap of 100 bp
* 151604 179979: contig of 28376 bp in length
* 179980 180079: gap of 100 bp
* 180080 215896: contig of 35817 bp in length.
FEATURES
    source
        1. 215896
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="15"
            /map="15"
            /clone="RP11-465J17"
            /clone_lib="RPCI-11 Human Male BAC"
        1. 1011
            /note="assembly_fragment"
        1112. 2633
            /note="assembly_fragment"
        2734. 4093
            /note="assembly_fragment"
        4194. 5298
            /note="assembly_fragment"
        5399. 7088
            /note="assembly_fragment"
        7189. 7889
            /note="assembly_fragment"
            clone_end:T7
            vector_side:right"
        7990. 10304
            /note="assembly_fragment"
        10405. 12955
            /note="assembly_fragment"
        13056. 16426
            /note="assembly_fragment"
        16527. 19178
            /note="assembly_fragment"
        19279. 23519
            /note="assembly_fragment"
        23620. 26853
            /note="assembly_fragment"
        26954. 31292
            /note="assembly_fragment"
        31393. 36401
            /note="assembly_fragment"
        36502. 41321
            /note="assembly_fragment"
        41422. 47286
            /note="assembly_fragment"
        47387. 54497
            /note="assembly_fragment"
        54598. 60757
            /note="assembly_fragment"
            clone_end:SP6
            vector_side:left"
        60858. 68004
            /note="assembly_fragment"
        68105. 77167
            /note="assembly_fragment"
        77268. 88567
            /note="assembly_fragment"
        88668. 101419
            /note="assembly_fragment"

```

```

/note="assembly_fragment"
101520..114909
/note="assembly_fragment"
115010..130182
/note="assembly_fragment"
130283..151503
/note="assembly_fragment"
151604..179979
/note="assembly_fragment"
180080..215896
/note="assembly_fragment"

ORIGIN
Query Match      68.7%; Score 20.6; DB 2; Length 215896;
Best Local Similarity 85.2%; Pred. No. 48;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      2  ACAATGTATGTCGGGTGACATCTATG 28
Db      4324  AAAACGTATGTCAGGTACATTTATG 4350

RESULT 18
AC073081/c
LOCUS      AC073081      118001 bp      DNA      linear      PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-310H21 from 2, complete sequence.
ACCESSION  AC073081
VERSION    AC073081.5 GI:16077041
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 118001)
            Suleston,J.E. and Waterston,R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
            99063792
            9847074
REFERENCE  2 (bases 1 to 118001)
            Nguyen,C., Kozlowicz,A., Boyer,E. and Bielicki,L.
            The sequence of Homo sapiens BAC clone RP11-310H21
            Unpublished (2001)
REFERENCE  3 (bases 1 to 118001)
            Waterston,R.H.
            Direct Submission
            Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  4 (bases 1 to 118001)
            Waterston,R.H.
            Direct Submission
            Submitted (12-OCT-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  5 (bases 1 to 118001)
            Waterston,R.
            Direct Submission
            Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Oct 12, 2001 this sequence version replaced gi:15528929.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_NH0310H21
            -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Ooegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-406M18, 2000 bp overlap. Actual end of this clone is at base position 118001 of RP11-310H21.

Data from AC092430 was used to finish this clone, AC073081. Polymorphisms have been identified between AC073081 and AC092430. A PCR only region exists between bases 98579 to 98669.

FEATURES

source	1..118001	/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="2"
		/map="2"
		/clone="RP11-310H21"
		/clone_lib="RPCI-11"
	1142..1164	
repeat_region		/rpt_family="(T)n"
repeat_region	2410..2463	
repeat_region		/rpt_family="(TCG)n"
repeat_region	4518..4543	
repeat_region		/rpt_family="AT-rich"
repeat_region	5397..6089	
repeat_region		/rpt_family="L1"
repeat_region	6215..7134	
repeat_region		/rpt_family="L1"
repeat_region	7149..7775	
repeat_region		/rpt_family="ERVL"
repeat_region	14542..14752	
repeat_region		/rpt_family="L1"
repeat_region	15940..16121	
repeat_region		/rpt_family="MIR"
repeat_region	16109..16235	
repeat_region		/rpt_family="L2"
repeat_region	16495..16638	
repeat_region		/rpt_family="(TA)n"
repeat_region	16689..16723	
repeat_region		/rpt_family="AT-rich"
repeat_region	16717..16820	
repeat_region		/rpt_family="MIR"
repeat_region	17385..17462	
repeat_region		/rpt_family="L2"
repeat_region	17944..18214	
repeat_region		/rpt_family="Alu"
repeat_region	21416..21480	
repeat_region		/rpt_family="L2"
repeat_region	21720..21816	
		/rpt_family="MIR"
		/rpt_family="MER1_type"
	22564..22642	
		/rpt_family="L2"
repeat_region	23859..23101	
		/rpt_family="MIR"
repeat_region	23509..23797	
		/rpt_family="Alu"
repeat_region	24221..24359	
		/rpt_family="MIR"
repeat_region	24410..24442	
		/rpt_family="AT-rich"
repeat_region	24962..25144	
		/rpt_family="ERVL"
repeat_region	25149..25203	
		/rpt_family="ERVL"
repeat_region	25189..25369	
		/rpt_family="(TA)n"
repeat_region	25381..25433	
		/rpt_family="(TA)n"
repeat_region	25429..25927	
		/rpt_family="ERVL"
repeat_region	25955..26078	
		/rpt_family="ERVL"
repeat_region	26220..27199	
		/rpt_family="ERVL"
repeat_region	27200..27490	
		/rpt_family="Alu"
repeat_region	27484..27521	
		/rpt_family="(GGAA)n"
repeat_region	27553..27610	
		/rpt_family="(GAAA)n"
repeat_region	27625..28477	
		/rpt_family="ERVL"
repeat_region	28478..28833	
		/rpt_family="MaLR"
repeat_region	28834..29073	
		/rpt_family="ERVL"
repeat_region	29093..30133	
		/rpt_family="ERVL"
repeat_region	30199..30624	
		/rpt_family="ERVL"
repeat_region	31533..31758	
		/rpt_family="L1"
repeat_region	32215..32616	
		/rpt_family="MaLR"
repeat_region	33073..33097	
		/rpt_family="(T)n"
repeat_region	33382..34641	
		/rpt_family="L1"
repeat_region	33653..33684	
		/rpt_family="A-rich"
repeat_region	34087..34142	
		/rpt_family="GA-rich"
repeat_region	34230..34254	
		/rpt_family="AT-rich"
repeat_region	34774..35137	
		/rpt_family="MaLR"
repeat_region	35170..35535	
		/rpt_family="L2"
repeat_region	35536..35822	
		/rpt_family="Alu"
repeat_region	35823..35871	
		/rpt_family="L2"
repeat_region	35972..36185	
		/rpt_family="Alu"
repeat_region	36186..36649	
		/rpt_family="L2"
repeat_region	36758..36780	
		/rpt_family="(TTCA)n"
repeat_region	36766..36860	
		/rpt_family="L2"
repeat_region	39431..39539	
		/rpt_family="MIR"


```

repeat_region 41473. 41782
/rpt_family="L1"
repeat_region 42017. 42054
/rpt_family="AT_rich"
repeat_region 42553. 42856

Query Match      58.0%; Score 20.4; DB 9; Length 118001;
Best Local Similarity 80.0%; Pred. No. 58;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AACAAATGATGTCGGGTGATCATATGAC 30
|||||
Db 81502 AACAAATGACTGGCAGGTGTCAGCTATGAC 81473
|||||

RESULT 19
AC040996/c
LOCUS      193910 bp      DNA      linear      HTG 28-MAY-2000
DEFINITION Homo sapiens chromosome 4 clone RP11-347K3 map 4, WORKING DRAFT
SEQUENCE   28 unordered pieces.
AC040996
VERSION    AC040996.2 GI:8099895
KEYWORDS   HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 193910)
AUTHORS   Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boglelavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On May 28, 2000 this sequence version replaced gi:7534198.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -----
            Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence.submissions@genome.wi.mit.edu
            -----
            Project Information
            -----
            Center project name: L9510
            Center clone name: 347_K_3
            -----
            Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 179519 bases at least Q40
            -----
            FEATURES
            source
            1. 193910
            /organism="Homo sapiens"
            Location/Qualifiers
            1164: contig of 1164 bp in length
            1165: gap of 100 bp
            1264: gap of 100 bp
            1265: contig of 347 bp in length
            1611: gap of 100 bp
            1711: gap of 100 bp
            1712: contig of 1468 bp in length
            3179: gap of 100 bp
            3279: gap of 100 bp
            3280: contig of 2378 bp in length
            5657: gap of 100 bp
            5658: contig of 1773 bp in length
            5758: gap of 100 bp
            7331: gap of 100 bp
            7332: contig of 1895 bp in length
            7631: gap of 100 bp
            9226: gap of 100 bp
            9226: contig of 2664 bp in length
            12289: gap of 100 bp
            12389: gap of 100 bp
            12390: contig of 2564 bp in length
            14954: gap of 100 bp
            15054: contig of 4431 bp in length
            19485: gap of 100 bp
            19485: contig of 3160 bp in length
            22745: gap of 100 bp
            22845: contig of 4964 bp in length
            27809: gap of 100 bp
            33103: contig of 5195 bp in length
            33203: gap of 100 bp
            33204: contig of 6207 bp in length
            39411: gap of 100 bp
            39510: gap of 100 bp
            45603: contig of 6093 bp in length
            45703: gap of 100 bp
            49406: contig of 3703 bp in length
            49406: gap of 100 bp
            56994: contig of 7488 bp in length
            57094: gap of 100 bp
            63379: contig of 6285 bp in length
            63479: gap of 100 bp
            71011: contig of 7532 bp in length
            71111: gap of 100 bp
            76827: contig of 5716 bp in length
            76927: gap of 100 bp
            84775: contig of 7848 bp in length
            84776: gap of 100 bp
            84776: contig of 5846 bp in length
            90721: gap of 100 bp
            90722: contig of 8063 bp in length
            90822: gap of 100 bp
            98885: contig of 5495 bp in length
            104480: gap of 100 bp
            104480: contig of 10719 bp in length
            115299: gap of 100 bp
            115299: contig of 12660 bp in length
            128059: gap of 100 bp
            128159: contig of 14766 bp in length
            142925: gap of 100 bp
            143025: contig of 20371 bp in length
            163396: gap of 100 bp
            163396: contig of 30415 bp in length.
            163496: contig of 30415 bp in length.
            -----
            Location/Qualifiers
            1. 193910
            /organism="Homo sapiens"

```

```

/misc_feature      /mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-347K3"
/clone_lib="RPC1-11 Human Male BAC"
1. .1164
/note="assembly_fragment"
1265. .1611
/note="assembly_fragment
clone_end:T7
vector_side:right"
1712. .3179
/note="assembly_fragment"
3280. .5657
/note="assembly_fragment"
5758. .7530
/note="assembly_fragment"
7631. .9525
/note="assembly_fragment"
9626. .12289
/note="assembly_fragment"
12390. .14953
/note="assembly_fragment"
15054. .19484
/note="assembly_fragment"
19585. .22744
/note="assembly_fragment"
22845. .27808
/note="assembly_fragment"
27909. .33103
/note="assembly_fragment"
33204. .39410
/note="assembly_fragment"
39511. .45603
/note="assembly_fragment"
45704. .49406
/note="assembly_fragment"
49507. .56994
/note="assembly_fragment"
57095. .63379
/note="assembly_fragment"
63480. .71011
/note="assembly_fragment"
71112. .76827
/note="assembly_fragment"
76928. .84775
/note="assembly_fragment"
84876. .90721
/note="assembly_fragment"
90822. .98884
/note="assembly_fragment"
98985. .104479
/note="assembly_fragment"
104580. .115298
/note="assembly_fragment"
115399. .128058
/note="assembly_fragment"
128159. .142924
/note="assembly_fragment"
143025. .163395
/note="assembly_fragment
clone_end:SP6
vector_side:left"
163496. .193910
/note="assembly_fragment"

```

ORIGIN

```

Query Match      68.0%; Score 20.4; DB 2; Length 193910;
Best Local Similarity 80.0%; Pred. No. 61;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 AACAAATGATGTCGGGTGATCATCTATGAC 30

```

```

Db      38306 AACAAATGACTGGCAGGTGTCCAGCTATGAC 38277

```

RESULT 20

```

AC109822/c
LOCUS

```

```

DEFINITION

```

```

AC109822
ACCESSION

```

```

AC109822.5 GI:21281638
VERSION

```

```

HYG.
KEYWORDS

```

```

SOURCE

```

```

ORGANISM

```

```

Homo sapiens
Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

1 (bases 1 to 194235)
AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

MEDLINE

```

```

99063792
PUBMED

```

```

2 (bases 1 to 194235)
REFERENCE

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

Unpublished (2001)
REFERENCE

```

```

3 (bases 1 to 194235)
AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

Submitted (07-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

4 (bases 1 to 194235)
REFERENCE

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

Submitted (06-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

5 (bases 1 to 194235)
REFERENCE

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

Submitted (31-MAY-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

6 (bases 1 to 194235)
REFERENCE

```

```

AUTHORS

```

```

TITLE

```

```

JOURNAL

```

```

Submitted (12-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 31, 2002 this sequence version replaced gi:19172920.

```

```

COMMENT

```

```

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
Summary Statistics
Center project name: H_NH0633J12

```

```

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

```

```

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

```

```

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.

```

McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Ooeagawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-742M10, 2000 bp overlap. Actual end is at base position 32235 of RP11-742M10.

Polymorphisms have been identified between AC092430 and AC109822. Data from AC073081 was used to finish this clone, AC109822.

FEATURES

source

1. 194235
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-633J12"
/clone_lib="RPCI-11"
1266..1374
/rpt_family="MIR"
3308..3617
/rpt_family="L1"
3852..3889
/rpt_family="AT_rich"
4388..4695
/rpt_family="Alu"
5730..5903
/rpt_family="L1"
6317..6415
/rpt_family="MER1_type"
6638..6660
/rpt_family="AT_rich"
6951..6982
/rpt_family="AT_rich"
8353..8546
/rpt_family="MIR"
8914..8940
/rpt_family="AT_rich"
9201..9479
/rpt_family="Alu"
9645..9775
/rpt_family="L2"
11441..11667
/rpt_family="L1"
11709..11859
/rpt_family="L1"
11860..12156
/rpt_family="Alu"
12157..12591
/rpt_family="L1"
12880..13015
/rpt_family="L1"
13760..14284
/rpt_family="MaLR"
14842..15205
/rpt_family="MaLR"
15592..15932
/rpt_family="MER1_type"
16611..16836
/rpt_family="MIR"
17542..17710
/rpt_family="MIR"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region 19658..19823
/rpt_family="MER1_type"
repeat_region 20081..20158
/rpt_family="L2"
21181..21411
/rpt_family="MaLR"
21636..21824
/rpt_family="MER1_type"
22076..22793
/rpt_family="L1"
24725..25022
/rpt_family="Alu"
25094..25431
/rpt_family="(TA)n"
25432..25462
/rpt_family="AT_rich"
25463..25958
/rpt_family="(TA)n"
25970..26137
/rpt_family="(TA)n"
26149..26313
/rpt_family="(TA)n"
26332..26625
/rpt_family="(TATATG)n"
26645..26819
/rpt_family="(TATATG)n"
26785..27197
/rpt_family="(TA)n"
27198..27248
/rpt_family="AT_rich"
27377..27411
/rpt_family="(TG)n"
27629..27823
/rpt_family="ERV1"
27824..28137
/rpt_family="Alu"
28138..28266
/rpt_family="ERV1"
28286..28459
/rpt_family="MIR"
28826..28858
/rpt_family="AT_rich"
28925..29380
/rpt_family="ERV1"
30224..30296
/rpt_family="L2"
30391..31169
/rpt_family="ERV1"
33587..33869
/rpt_family="Alu"
34870..34901
/rpt_family="AT_rich"
35012..35379
/rpt_family="ERV1"
35381..35423
/rpt_family="(CATATA)n"
35462..35498
/rpt_family="(TG)n"
35587..35873
/rpt_family="Alu"

Query Match 58.0%; Score 20.4; DB 9; Length 194235;
Best Local Similarity 80.0%; Pred. No. 61;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AACAAATGATGTCGGGTGATCATCTATGAC 30

Db 43306 AACAAATGACTGGCAGGTGTCAGCTATGAC 43277

RESULT 21

AC092430/c

LOCUS

AC092430

288385 bp

DNA

linear

HTG 05-SEP-2001

```

DEFINITION Homo sapiens chromosome 2 clone RP11-71H2, WORKING DRAFT SEQUENCE,
12 unordered pieces.
ACCESSION AC092430
VERSION AC092430.1 GI:14595952
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 288385)
Waterston,R.H.
Direct Submission
Submitted (04-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0071H02
----- Summary Statistics -----
Sequencing vector: M13, 52%
Sequencing vector: plasmid, 48%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 278642 bases at least Q40
Consensus quality: 282293 bases at least Q30
Consensus quality: 284168 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2137: contig of 2137 bp in length
* 2138 2237: gap of unknown length
* 2238 6019: contig of 3782 bp in length
* 6020 6119: gap of unknown length
* 6120 11179: contig of 5060 bp in length
* 11180 11279: gap of unknown length
* 11280 19638: contig of 8359 bp in length
* 19639 19738: gap of unknown length
* 19739 25335: contig of 5597 bp in length
* 25336 25435: gap of unknown length
* 25436 36165: contig of 10730 bp in length
* 36166 36265: gap of unknown length
* 36266 48376: contig of 12110 bp in length
* 48376 48476: gap of unknown length
* 48476 75051: contig of 26575 bp in length
* 75051 106333: contig of 31183 bp in length
* 106334 106433: gap of unknown length
* 106434 132723: contig of 26290 bp in length
* 132724 132824: gap of unknown length
* 132824 168406: contig of 35583 bp in length
* 168407 168506: gap of unknown length
* 168507 288385: contig of 119879 bp in length.
Location/Qualifiers
1..288385
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

/chrromosome="2"
/clone="RP11-71H2"
1..2137
/notes="assembly_name:Contig26"
2238..6019
/notes="assembly_name:Contig27"
6120..11179
/notes="assembly_name:Contig28"
clone_end:77
vector_side:left"
11280..19638
/notes="assembly_name:Contig29"
19739..25335
/notes="assembly_name:Contig30"
25436..36165
/notes="assembly_name:Contig31"
36266..48375
/notes="assembly_name:Contig32"
48476..75050
/notes="assembly_name:Contig33"
75151..106333
/notes="assembly_name:Contig34"
106434..132723
/notes="assembly_name:Contig35"
132824..168406
/notes="assembly_name:Contig36"
168507..288385
/notes="assembly_name:Contig37"
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 68.0%; Score 20.4; DB 2; Length 288385;
Best Local Similarity 80.0%; Pred No. 63;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCCGGTGATCATCTATGAC 30
||||| ||||| ||||| ||||| |||||
Db 253932 AACAAATGACTGGCAGGTGTCAGCTATGAC 253903

RESULT 22
AC008034 121101 bp DNA linear HTG 21-SEP-2002
LOCUS Homo sapiens chromosome 3 clone CTB-134A23, WORKING DRAFT SEQUENCE,
DEFINITION 4 unordered pieces.
ACCESSION AC008034
VERSION AC008034.30 GI:22138152
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEPIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoggues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichte, O., Lieu, C., Liu, J., Liu, W., Loulesed, H., Lozada, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mahoney, E., McLeod, M. P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, S., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviaden, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sison, I., Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----
Center project name: HMHO
Center clone name: CTB-134A23
----- Summary Statistics -----
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 3% of reads
Chemistry: Dye-terminator Big Dye; 6% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 119470 bases at least Q40
Consensus quality: 120027 bases at least Q30
Consensus quality: 120436 bases at least Q20
Estimated insert size: 121022; sum-of-contigs estimation
Quality coverage: 11x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html). -----
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2831: contig of 2831 bp in length
* 2832 2931: gap of unknown length
* 2932 27219: contig of 24288 bp in length
* 27220 27319: gap of unknown length
* 27320 81157: contig of 53836 bp in length
* 81158 81257: gap of unknown length

* 81258 121101: contig of 39844 bp in length.

FEATURES
source
Location/Qualifiers
1. 121101
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="CTB-134A23"

ORIGIN
Query Match 67.3%; Score 20.2; DB 2; Length 121101;
Best Local Similarity 88.0%; Pred No. 74;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ATCTATGTCGGGTACATCTATGA 29
|||||
Db 93892 ATCTATGTCGAAGTGATCATCTGTGA 93916
|||||

RESULT 23
AC068315
LOCUS
DEFINITION
Homo sapiens chromosome 3 clone RP11-224P21 map 3p, WORKING DRAFT
SEQUENCE, 28 unordered pieces.
AC068315
VERSION
AC068315.4 GI:10312236
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 155313)
Bao, J., Bao, Q., Kang, N., Li, C., Li, G., Li, J., Li, L., Huang, F., Jin, Y., Kang, N., Li, C., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, S., Sun, M., Sun, Y., Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yang, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Yu, J. and Yang, H.
Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 155313)
Bao, W., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.
Direct Submission
Submitted (02-MAY-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P. R. China
On Sep 27, 2000 this sequence version replaced gi:8101156.

-----Genome Center
Center: Beijing Center
Center code: Beijing
Website: <http://hgsc.igtp.ac.cn>
<http://www.genomics.org.cn>
Contact: hgsc@igtp.ac.cn

----- Project Information -----
Center project name: 1# project
Center clone name: RP11-224P21
----- Summary Statistics -----
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 154843 bases at least Q40
Consensus quality: 165011 bases at least Q30
Consensus quality: 168985 bases at least Q20

Insert size: 139544; sum-of-contigs
Quality coverage: 5.36x in Q20 bases;sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
* 1139: contig of 1139 bp in length
* 1233: gap of unknown length
* 1440
* 2448: contig of 1209 bp in length
* 2449
* 2548: gap of unknown length
* 2549
* 3935: contig of 1387 bp in length
* 3936
* 4035: gap of unknown length
* 4036
* 6131: contig of 2096 bp in length
* 6132
* 6231: gap of unknown length
* 6232
* 7423: contig of 1198 bp in length
* 7430
* 7529: gap of unknown length
* 7530
* 9054: contig of 1525 bp in length
* 9055
* 9154: gap of unknown length
* 10950: contig of 1796 bp in length
* 10951
* 11050: gap of unknown length
* 12637: contig of 1587 bp in length
* 12638
* 12737: gap of unknown length
* 15121: contig of 2384 bp in length
* 15122
* 15221: gap of unknown length
* 17220: contig of 1999 bp in length
* 17221
* 17320: gap of unknown length
* 17321
* 20436: contig of 3116 bp in length
* 20437
* 20536: gap of unknown length
* 23761: contig of 3225 bp in length
* 23762
* 23861: gap of unknown length
* 26190: contig of 2329 bp in length
* 26191
* 26290: gap of unknown length
* 26291
* 28871: contig of 2581 bp in length
* 28872
* 28971: gap of unknown length
* 3062: contig of 4091 bp in length
* 3063
* 3162: gap of unknown length
* 3163
* 37033: contig of 3871 bp in length
* 37133: gap of unknown length
* 37134
* 43465: contig of 6332 bp in length
* 43466
* 43565: gap of unknown length
* 48805: contig of 5240 bp in length
* 48806
* 48905: gap of unknown length
* 56158: contig of 7253 bp in length
* 56159
* 62458: gap of unknown length
* 62459
* 63409: contig of 7151 bp in length
* 63410
* 63509: gap of unknown length
* 72533: contig of 9030 bp in length
* 72540
* 72633: gap of unknown length
* 83389: contig of 10750 bp in length
* 83390
* 83489: gap of unknown length
* 91426: contig of 7937 bp in length
* 91427
* 91526: gap of unknown length
* 95009: contig of 7983 bp in length
* 95010
* 99609: gap of unknown length
* 108238: contig of 8629 bp in length
* 108239
* 108338: gap of unknown length
* 108339
* 119079: contig of 10741 bp in length
* 119080
* 119179: gap of unknown length
* 119180
* 135925: contig of 16746 bp in length
* 135926
* 136025: gap of unknown length
* 136026
* 155313: contig of 19288 bp in length.

FEATURES
source

1. 155313
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"

misc_feature
1. 1139
/note="assembly_name:Contig32"
1240. 2448
/note="assembly_name:Contig33"
2549. 3935
/note="assembly_name:Contig34"
4036. 6131
/note="assembly_name:Contig35"
6232. 7429
/note="assembly_name:Contig36"
7530. 9054
/note="assembly_name:Contig37"
9155. 10950
/note="assembly_name:Contig38"
11051. 12637
/note="assembly_name:Contig39"
12738. 15121
/note="assembly_name:Contig40"
15222. 17220
/note="assembly_name:Contig41"
17321. 20436
/note="assembly_name:Contig42"
20537. 23761
/note="assembly_name:Contig43"
23862. 26190
/note="assembly_name:Contig44"
26291. 28871
/note="assembly_name:Contig45"
28972. 33062
/note="assembly_name:Contig46"
33163. 37033
/note="assembly_name:Contig47"
37134. 43465
/note="assembly_name:Contig48"
43566. 48805
/note="assembly_name:Contig49"
48906. 56158
/note="assembly_name:Contig50"
56259. 63409
/note="assembly_name:Contig51"
63510. 72539
/note="assembly_name:Contig52"
clone_end:SP6
vector_side:right"
72640. 83389
/note="assembly_name:Contig53"
83490. 91426
/note="assembly_name:Contig54"
91527. 99509
/note="assembly_name:Contig55"
99610. 108238
/note="assembly_name:Contig56"
108339. 119079
/note="assembly_name:Contig57"
119180. 135925
/note="assembly_name:Contig58"
136026. 155313
/note="assembly_name:Contig59"

ORIGIN

Query Match 67.3%; Score 20.2; DB 2; Length 155313;
Best Local Similarity 88.0%; Pred.No.75;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTACATCTATGA 29
|||||
Db 42000 ATGTATGTCGAAGTGTACATCTGTGA 42024
|||||

RESULT 24
AC066599
LOCUS

AC066599 157007 bp DNA linear HTG 19-SEP-2000

```

DEFINITION Homo sapiens chromosome 3 clone RP11-596J9 map 3p, WORKING DRAFT
SEQUENCE, 14 unordered pieces.
ACCESSION AC066599
VERSION AC066599.3 GI:10190763
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157007)
Bao,J., Bao,Q., Wang,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,Y.,
Tao,R., Wang,H., Wang,J., Wang,L., Wang,L., Wang,R.,
Wang,X., Wang,X., Wang,Y., Wu,D., Xie,F., Xuan,Z., Xue,Y.,
Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
Zhang,Z., Zhu,B., Yu,J. and Yang,H.
Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 157007)
Wang,J., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
and Yang,H.
Direct Submission
Submitted (25-APR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On Sep 19, 2000 this sequence version replaced gi:8101249.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hg.cigtp.ac.cn
http://www.genomics.org.cn
Contact:hgc@cigtp.ac.cn
----- Project Information
Center project name:11 project
Center clone name: RP11-596J9
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 55% of reads
Assembly program: Big Dye; 45% of reads
Consensus quality: 150014 bases at least Q40
Consensus quality: 153515 bases at least Q30
Consensus quality: 155718 bases at least Q20
Insert size: 149003; sum-of-contigs
Quality coverage: 4.98x in Q20 bases,sum-of-contigs
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1641: contig of 1641 bp in length
* 1642: gap of unknown length
* 1742: contig of 4070 bp in length
* 5812: gap of unknown length
* 5912: contig of 1951 bp in length
* 7862: gap of unknown length
* 7863: gap of unknown length
* 11614: contig of 3652 bp in length
* 11715: gap of unknown length
* 11715: contig of 3914 bp in length
* 15629: gap of unknown length
* 15729: contig of 6471 bp in length

```

```

* 22200 22299: gap of unknown length
* 22300 32786: contig of 10487 bp in length
* 32787 32886: gap of unknown length
* 32887 39504: contig of 6618 bp in length
* 39505 39604: gap of unknown length
* 39605 48115: contig of 8511 bp in length
* 48116 48215: gap of unknown length
* 48216 57788: contig of 9573 bp in length
* 57789 57888: gap of unknown length
* 57889 69171: contig of 11283 bp in length
* 69172 69272: gap of unknown length
* 69272 86305: contig of 17034 bp in length
* 86306 86405: gap of unknown length
* 86406 117887: contig of 31482 bp in length
* 117888 117987: gap of unknown length
* 117988 157007: contig of 39020 bp in length.
FEATURES
Location/Qualifiers
1..157007
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-596J9"
misc_feature 1..1641
/notes="assembly_name:Contig5"
misc_feature 1742..5811
/notes="assembly_name:Contig6"
misc_feature 5912..7862
/notes="assembly_name:Contig7"
misc_feature 7863..11614
/notes="assembly_name:Contig8"
misc_feature 11715..15628
/notes="assembly_name:Contig9"
misc_feature 15729..22199
/notes="assembly_name:Contig10"
misc_feature 22300..32786
/notes="assembly_name:Contig11"
misc_feature 32887..39504
/notes="assembly_name:Contig12"
misc_feature 39605..48115
/notes="assembly_name:Contig13"
misc_feature 48216..57788
/notes="assembly_name:Contig14"
misc_feature 57889..69171
/notes="assembly_name:Contig15"
misc_feature 69272..86305
/notes="assembly_name:Contig16"
misc_feature 86406..117887
/notes="assembly_name:Contig17"
misc_feature 117988..157007
/notes="assembly_name:Contig18"
ORIGIN

```

```

Query Match 67.3%; Score 20.2; DB 2; Length 157007;
Best Local Similarity 88.0%; Pred. No. 75;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ATGTATGTCGGGTGATCATCTATGA 29
|||||
DB 150679 ATGTATGTCAGTGATCATCTGTA 150703

RESULT 25
AC021996/c
LOCUS AC021996 161903 bp DNA linear HTG 03-FEB-2000
DEFINITION Homo sapiens chromosome 3p clone RP11-1087020, WORKING DRAFT
SEQUENCE, 39 unordered pieces.
ACCESSION AC021996
VERSION AC021996.1 GI:6742988
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 161903)
Zeng, Y., Hu, S., Dong, W., Zhang, X., Wang, J., Wang, X., Zhang, Y.,
Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X.,
Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D.,
Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X.,
Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X.,
Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

TITLE
JOURNAL

Chromosome 3p genomic sequence

REFERENCE
AUTHORS

2 (bases 1 to 161903)
Zeng, Y., Hu, S., Dong, W., Zhang, X., Wang, J., Wang, X., Zhang, Y.,
Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X.,
Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D.,
Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X.,
Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X.,
Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

TITLE
JOURNAL

Submitted (24-JAN-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2162: contig of 2162 bp in length

* 2163 gap of unknown length

* 3555 contig of 1392 bp in length

* 5164 gap of unknown length

* 6572 contig of 1609 bp in length

* 9023 gap of unknown length

* 10618 contig of 1408 bp in length

* 12744 gap of unknown length

* 14961 contig of 2451 bp in length

* 16934 gap of unknown length

* 18727 contig of 1595 bp in length

* 20269 gap of unknown length

* 22025 contig of 2126 bp in length

* 24218 gap of unknown length

* 27013 contig of 1793 bp in length

* 29738 gap of unknown length

* 31315 contig of 2217 bp in length

* 34025 gap of unknown length

* 37281 contig of 1542 bp in length

* 40668 gap of unknown length

* 43988 contig of 2710 bp in length

* 47757 gap of unknown length

* 51701 contig of 3256 bp in length

* 51701 contig of 3387 bp in length

* 51701 contig of 3320 bp in length

* 51701 contig of 3769 bp in length

* 51701 contig of 3945 bp in length

* 51702 gap of unknown length

* 56620 contig of 4918 bp in length

* 60241 gap of unknown length

* 65976 contig of 3621 bp in length

* 70349 gap of unknown length

* 76647 contig of 5735 bp in length

* 82486 gap of unknown length

* 86969 contig of 4373 bp in length

* 93335 gap of unknown length

* 97637 contig of 6298 bp in length

* 102998 gap of unknown length

* 109895 contig of 5839 bp in length

* 116212 gap of unknown length

* 124568 contig of 4483 bp in length

* 133484 gap of unknown length

* 140452 contig of 6366 bp in length

* 149744 gap of unknown length

* 161903 contig of 4303 bp in length

* 161903 contig of 5361 bp in length

* 161903 contig of 6896 bp in length

* 161903 contig of 6317 bp in length

* 161903 contig of 8357 bp in length

* 161903 contig of 8915 bp in length

* 161903 contig of 6968 bp in length

* 161903 contig of 9292 bp in length

* 161903 contig of 12160 bp in length

FEATURES
Source

1. 161903
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3p"
/clone="RP11-1087020"

ORIGIN

Query Match 67.3% Score 20.2; DB 2; Length 161903;
Best Local Similarity 88.0%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTACATCTATGA 29

Db 1257 ATGTATGTCGAAGTGATACATCTGTGA 1233

Search completed: October 28, 2005, 11:16:55

Job time : 1484 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 05:33:59 ; Search time 234 Seconds
(without alignments)
758.941 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgatgtccggtgacatcatgac 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001s:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003s:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	AAA37981	Aaa37981 PCR prime
2	30	100.0	2048	AAA37962	Aaa37962 Plasmid #
3	30	100.0	2048	AAA37963	Aaa37963 Plasmid #
4	30	100.0	2052	AAA37959	Aaa37959 Carrot pr
5	30	100.0	2052	AAA37961	Aaa37961 Carrot pr
6	30	100.0	2056	AAA37964	Aaa37964 Plasmid #
7	23.6	78.7	2042	AAV15144	AAV15144 New promo
8	19.6	65.3	2931	ADN73150	Adn73150 Thale cre
9	19.4	64.7	110000	11 ACN44932_0	Acn44932 Mouse gen
10	19.2	64.0	42772	11 ACN45116	Acn45116 Mouse gen
11	19	63.3	512	13 ACN51227	Acn51227 Cotton an
12	19	63.3	2700	10 ADB53615	Adb53615 Primary r
13	19	63.3	35042	3 AAA81454	Aaa81454 N. mening
14	19	63.3	110000	3 AAA81489_1	Continuation (2 of
15	19	63.3	236303	4 AAS11614	Aas11614 Human gen
16	19	63.3	349880	3 AAF21611	Aaf21611 Neisseria
17	18.8	62.7	1630	6 AAS62657	Aas62657 cDNA sequ
18	18.6	62.0	1569	3 AAA30829	Aaa30829 Zebrafish
19	18.6	62.0	1671	3 AAA48445	Aaa48445 Zebrafish
20	18.6	62.0	1671	10 ADC42329	Adc42329 cDNA enco

21	18.6	62.0	2152	3 AAA48446	Aaa48446 Zebrafish
22	18.6	62.0	2152	10 ADC42306	Adc42306 cDNA enco
23	18.6	62.0	2152	10 ADH61060	Adh61060 Zebrafish
24	18.6	62.0	90442	9 ADA03077	Ada03077 Mouse mCG
25	18.6	62.0	90442	9 ADA66361	Ada66361 Mouse mCG
26	18.6	62.0	90442	10 ADB72815	Adb72815 Mouse mCG
27	18.6	62.0	90442	10 ADC26997	Adc26997 Mouse car
28	18.6	62.0	90442	11 ADL27155	Adl27155 Mouse gen
29	18.4	61.3	2501	13 ADR73232	Adr73232 Thale cre
30	18.4	61.3	3177	4 ABL28637	AbL28637 Drosophil
31	18.4	61.3	4354	4 ABL11038	AbL11038 Drosophil
32	18.4	61.3	5460	4 ABL28636	AbL28636 Drosophil
33	18.4	61.3	188017	11 ACN45148	Acn45148 Mouse gen
34	18.2	60.7	4919	9 ABL15204	AbL15204 Drosophil
35	18.2	60.7	40050	9 ADA02585	Ada02585 Mouse ics
36	18.2	60.7	40050	10 ADR72323	Adr72323 Mouse ics
37	18.2	60.7	40050	10 ADR58333	Adr58333 Mouse ics
38	18.2	60.7	310122	13 ABD32533	Abd32533 Mouse can
39	18	60.0	287	12 ADL11418	AdL11418 Cat flea
40	18	60.0	287	12 ADL11392	AdL11392 Cat flea
41	18	60.0	386	12 ADL11486	AdL11486 Cat flea
42	18	60.0	401	4 AAK96285	Aak96285 Human neu
43	18	60.0	401	4 AAK97778	Aak97778 Human neu
44	18	60.0	401	6 ABT01055	Abt01055 Human neu
45	18	60.0	401	6 ABT02548	Abt02548 Human neu
46	18	60.0	448	12 ADL10839	AdL10839 Cat flea
47	18	60.0	83709	12 ADQ19964	Adq19964 Human sof
48	18	60.0	110000	4 AAK95240_10	Continuation (11 o
49	18	60.0	110000	4 AAK95240_11	Continuation (12 o
50	18	60.0	110000	4 AAK96733_10	Continuation (11 o
51	18	60.0	110000	4 AAK96733_11	Continuation (12 o
52	18	60.0	110000	6 ABT00010_10	Continuation (11 o
53	18	60.0	110000	6 ABT00010_11	Continuation (12 o
54	18	60.0	110000	6 ABT01503_10	Continuation (11 o
55	18	60.0	110000	6 ABT01503_11	Continuation (12 o
56	18	60.0	110000	12 ADH77486_10	Continuation (11 o
57	18	60.0	110000	12 ADH77486_11	Continuation (12 o
58	17.8	59.3	377	4 AAL09572	Aal09572 Human bre
59	17.8	59.3	379	4 AAL17464	Aal17464 Human bre
60	17.8	59.3	428	11 ACN79772	Acn79772 Breast ca
61	17.8	59.3	440	5 AAS67070	Aas67070 DNA encod
62	17.8	59.3	500	9 ACH36742	Ach36742 Human end
63	17.8	59.3	598	4 AAL09479	Aal09479 Human bre
64	17.8	59.3	601	4 AAL17368	Aal17368 Human bre
65	17.8	59.3	662	11 ACN79679	Acn79679 Breast ca
66	17.8	59.3	881	11 ACN89461	Acn89461 Breast ca
67	17.8	59.3	908	5 AAD03340	Aad03340 Probe 8 f
68	17.8	59.3	908	10 ADB91926	Adb91926 Acyltrans
69	17.8	59.3	1425	3 AAC40897	Aac40897 Arabidops
70	17.8	59.3	4338	13 ADS47498	AdS47498 Bacterial
71	17.8	59.3	8893	4 ABL13945	AbL13945 Drosophil
72	17.8	59.3	9739	4 ABL20301	AbL20301 Drosophil
73	17.8	59.3	12979	4 ABL13944	AbL13944 Drosophil
74	17.8	59.3	14727	4 ABL20300	AbL20300 Drosophil
75	17.8	59.3	99014	6 ABN96931	Abn96931 Gene #342
76	17.8	59.3	121129	13 ABD33446	Abd33446 Murine ca
77	17.8	59.3	215126	12 ADQ97362	Adq97362 Mouse can
78	17.6	58.7	201	13 ADS37021	AdS37021 Human aut
79	17.6	58.7	810	3 AAC79975	Aac79975 Human sec
80	17.6	58.7	963	3 AAC47090	Aac47090 Arabidops
81	17.6	58.7	1761	3 AAC51203	Aac51203 Arabidops
82	17.6	58.7	2022	3 AAC50796	Aac50796 Arabidops
83	17.6	58.7	2535	13 ADR07452	Adr07452 Full leng
84	17.6	58.7	4470	12 ADJ39886	Adj39886 Plant CDN
85	17.6	58.7	4508	4 AAL03274	Aal03274 Human rep
86	17.6	58.7	5084	4 AAL03275	Aal03275 Human rep
87	17.6	58.7	5087	4 AAL03276	Aal03276 Human rep
88	17.6	58.7	42339	11 ACN44508	Acn44508 Mouse gen
89	17.6	58.7	94330	11 ACN44662	Acn44662 Human gen
90	17.6	58.7	321019	13 ADS36450	AdS36450 Human aut
91	17.6	58.7	329019	13 ABD32707	Abd32707 Human can
92	17.4	58.0	378	5 AAF67029	Aaf67029 Novel hum
93	17.4	58.0	412	10 ADF66363	Adf66363 Human mic

```

c 94 17.4 58.0 425 5 AAF66854
95 17.4 58.0 452 4 AAI83581
96 17.4 58.0 483 12 ADL11119
c 97 17.4 58.0 663 6 AAS61924
98 17.4 58.0 699 5 AAS67755
99 17.4 58.0 1731 5 ABV24103
100 17.4 58.0 1731 5 ABV29984

```

ALIGNMENTS

RESULT 1

```

AAA37981
ID AAA37981 standard; DNA; 30 BP.

```

```

XX AC AAA37981;

```

```

XX DT 18-AUG-2000 (first entry)

```

```

XX DE PCR primer S used for carrot promoter amplification.

```

```

XX KW Carrot; promoter; terminator; transgenic plant; breeding; fertility;
XX KW PCR primer; ss.

```

```

XX OS Daucus carota.

```

```

XX PN WO200020613-A1.

```

```

XX PD 13-APR-2000.

```

```

XX PF 28-SEP-1999; 99WO-JP005303.

```

```

XX PR 02-OCT-1998; 98JP-00281124.

```

```

XX PA (SUMO ) SUMITOMO CHEM CO LTD.

```

```

XX PI Nishikawa S, Oeda K;

```

```

XX DR WPI; 2000-303791/26.

```

```

XX PT New Plant promoters and terminators from Daucus carota L., useful in
XX PT plant breeding, for e.g. controlling fertilities of plants.

```

```

XX PS Example 6; Page 44; 81pp; English.

```

```

XX CC This sequence represents a PCR primer used to amplify a carrot promoter
XX CC sequence. The invention relates to plant promoters and terminators from
XX CC Daucus carota L. which are capable of expressing a gene of interest in
XX CC plants. The invention also includes a chimeric gene characterized in that
XX CC it comprises the promoter and a desired gene linked to each other in the
XX CC form capable of functioning. A method of producing a transformant
XX CC comprises introducing the promoter, the chimeric gene or a vector
XX CC comprising the promoter and a desired gene or terminator sequence into a
XX CC host cell. The plant promoters and terminators are useful in plant
XX CC breeding, for e.g. fertilities of plants may be controlled by expressing,
XX CC in the host cells, a sense or antisense gene of a male sterility related
XX CC gene such as S-locus-specific RNase gene

```

```

XX SQ Sequence 30 BP; 9 A; 6 C; 6 G; 9 T; 0 U; 0 Other;

```

```

Query Match 100.0%; Score 30; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AACAAATGTATGTCGGGTGTACATCTATGAC 30

```

```

Db 1 AACAAATGTATGTCGGGTGTACATCTATGAC 30

```

RESULT 2

```

AAA37962

```

```

ID AAA37962 standard; DNA; 2048 BP.

```

```

XX AC AAA37962;

```

```

XX DT 18-AUG-2000 (first entry)

```

```

XX DE Plasmid #1 DNA sequence used in mutation of promoter sequence.

```

```

XX KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

```

```

XX OS Daucus carota.

```

```

XX PN WO200020613-A1.

```

```

XX PD 13-APR-2000.

```

```

XX PF 28-SEP-1999; 99WO-JP005303.

```

```

XX PR 02-OCT-1998; 98JP-00281124.

```

```

XX PA (SUMO ) SUMITOMO CHEM CO LTD.

```

```

XX PI Nishikawa S, Oeda K;

```

```

XX DR WPI; 2000-303791/26.

```

```

XX PT New Plant promoters and terminators from Daucus carota L., useful in
XX PT plant breeding, for e.g. controlling fertilities of plants.

```

```

XX PS Example 8; Page 71-73; 81pp; English.

```

```

XX CC This sequence represents a plasmid sequence used in a method for
XX CC introducing a mutation into a carrot promoter. The invention relates to
XX CC plant promoters and terminators from Daucus carota L. which are capable
XX CC of expressing a gene of interest in plants. The invention also includes a
XX CC chimeric gene characterized in that it comprises the promoter and a
XX CC desired gene linked to each other in the form capable of functioning. A
XX CC method of producing a transformant comprises introducing the promoter,
XX CC the chimeric gene or a vector comprising the promoter and a desired gene
XX CC or terminator sequence into a host cell. The plant promoters and
XX CC terminators are useful in plant breeding, for e.g. fertilities of plants
XX CC may be controlled by expressing, in the host cells, a sense or antisense
XX CC gene of a male sterility related gene such as S-locus-specific RNase gene

```

```

XX SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

```

```

Query Match 100.0%; Score 30; DB 3; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.0004;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AACAAATGTATGTCGGGTGTACATCTATGAC 30

```

```

Db 1742 AACAAATGTATGTCGGGTGTACATCTATGAC 1771

```

RESULT 3

```

AAA37963

```

```

ID AAA37963 standard; DNA; 2048 BP.

```

```

XX AC AAA37963;

```

```

XX DT 18-AUG-2000 (first entry)

```

```

XX DE Plasmid #2 DNA sequence used in mutation of promoter sequence.

```

```

XX KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

```

```

XX OS Daucus carota.

```

```

XX PN WO200020613-A1.

```

```

XX PD 13-APR-2000.

```

```

XX PF 28-SEP-1999; 99WO-JP005303.

```

XX 02-OCT-1998; 98JP-00281124.
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX Nishikawa S, Oeda K;
 XX WPI; 2000-303791/26.
 XX New Plant promoters and terminators from Daucus carota L., useful in
 XX plant breeding, for e.g. controlling fertilities of plants.
 XX Example 8; Page 73-74; 81pp; English.
 XX This sequence represents a plasmid sequence used in a method for
 XX introducing a mutation into a carrot promoter. The invention relates to
 XX plant promoters and terminators from Daucus carota L. which are capable
 XX of expressing a gene of interest in plants. The invention also includes a
 XX chimeric gene characterized in that it comprises the promoter and a
 XX desired gene linked to each other in the form capable of functioning. A
 XX method of producing a transformant comprises introducing the promoter,
 XX the chimeric gene or a vector comprising the promoter and a desired gene
 XX or terminator sequence into a host cell. The plant promoters and
 XX terminators are useful in plant breeding, for e.g. fertilities of plants
 XX may be controlled by expressing, in the host cells, a sense or antisense
 XX gene of a male sterility related gene such as S-locus-specific RNase gene
 XX Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 30; DB 3; Length 2048;
 Best Local Similarity 100.0%; Pred. No. 0.0004;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
 Db 1742 AACAAATGATGTCGGGTGTACATCTATGAC 1771

RESULT 4
 AAA37959
 ID AAA37959 standard; DNA; 2052 BP.
 XX AAA37959;
 AC
 DT 18-AUG-2000 (first entry)
 XX Carrot promoter sequence #1.
 DE Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
 XX Daucus carota.
 OS
 PN WO2000020613-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 28-SEP-1999; 99WO-JP005303.
 XX
 PR 02-OCT-1998; 98JP-00281124.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 PI Nishikawa S, Oeda K;
 XX WPI; 2000-303791/26.
 XX New Plant promoters and terminators from Daucus carota L., useful in
 XX plant breeding, for e.g. controlling fertilities of plants.
 XX Claim 1; Page 69-70; 81pp; English.
 XX This sequence represents a carrot promoter. The invention relates to
 XX plant promoters and terminators from Daucus carota L. which are capable

CC of expressing a gene of interest in plants. The invention also includes a
 CC chimeric gene characterized in that it comprises the promoter and a
 CC desired gene linked to each other in the form capable of functioning. A
 CC method of producing a transformant comprises introducing the promoter,
 CC the chimeric gene or a vector comprising the promoter and a desired gene
 CC or terminator sequence into a host cell. The plant promoters and
 CC terminators are useful in plant breeding, for e.g. fertilities of plants
 CC may be controlled by expressing, in the host cells, a sense or antisense
 CC gene of a male sterility related gene such as S-locus-specific RNase gene
 XX Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 30; DB 3; Length 2052;
 Best Local Similarity 100.0%; Pred. No. 0.00041;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
 Db 1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775

RESULT 5
 AAA37961
 ID AAA37961 standard; DNA; 2052 BP.
 XX AAA37961;
 AC
 DT 18-AUG-2000 (first entry)
 XX Carrot promoter sequence #2.
 DE Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
 XX Daucus carota.
 OS
 PN WO2000020613-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 28-SEP-1999; 99WO-JP005303.
 XX
 PR 02-OCT-1998; 98JP-00281124.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 PI Nishikawa S, Oeda K;
 XX WPI; 2000-303791/26.
 XX New Plant promoters and terminators from Daucus carota L., useful in
 XX plant breeding, for e.g. controlling fertilities of plants.
 XX Claim 1; Page 78-79; 81pp; English.
 XX This sequence represents a carrot promoter. The invention relates to
 XX plant promoters and terminators from Daucus carota L. which are capable
 XX of expressing a gene of interest in plants. The invention also includes a
 XX chimeric gene characterized in that it comprises the promoter and a
 XX desired gene linked to each other in the form capable of functioning. A
 XX method of producing a transformant comprises introducing the promoter,
 XX the chimeric gene or a vector comprising the promoter and a desired gene
 XX or terminator sequence into a host cell. The plant promoters and
 XX terminators are useful in plant breeding, for e.g. fertilities of plants
 XX may be controlled by expressing, in the host cells, a sense or antisense
 XX gene of a male sterility related gene such as S-locus-specific RNase gene
 XX Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 30; DB 3; Length 2052;
 Best Local Similarity 100.0%; Pred. No. 0.00041;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30

```

Db      1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775
|||||
RESULT 6
AAV15144
ID AAV15144 standard; DNA; 2056 BP.
XX
XX
AC AAV15144;
XX
XX DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
XX Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
XX WO2000020613-A1.
XX
XX PD 13-APR-2000.
XX
XX PF 28-SEP-1999; 99WO-JP005303.
XX
XX PR 02-OCT-1998; 98JP-00281124.
XX
XX PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX PI Nishikawa S, Oeda K;
XX
XX WIPI; 2000-303791/26.
XX
XX PT New plant promoters and terminators from Daucus carota L., useful in
XX plant breeding, for e.g. controlling fertilities of plants.
XX
XX PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter, a
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
XX SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;
Query Match 100.0%; Score 30; DB 3; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AACAAATGATGTCGGGTGTACATCTATGAC 30
|||||
Db      1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775
|||||
RESULT 7
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX
XX
AC AAV15144;
XX
XX DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
XX Promoter; root; carrot; Kuroda Gosun; root-specific expression;
XX plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW

```

```

KW
XX
OS ss.
XX Daucus carota.
XX
XX PN EP824150-A2.
XX
XX PD 18-FEB-1998.
XX
XX PF 12-AUG-1997; 97EP-00113923.
XX
XX PR 12-AUG-1996; 96JP-00212680.
XX
XX PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX PI Torikai S, Oeda K;
XX
XX WIPI; 1998-122310/12.
XX
XX PT New carrot root gene, promoter and terminator - useful in genetic
XX engineering for directing root-specific gene expression.
XX
XX PS Claim 2; Page 15-16; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
XX SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
Query Match 78.7%; Score 23.6; DB 2; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.5;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AACAAATGATGTCGGGTGTACATCTATGAC 30
|||||
Db      1737 AACAAATGATGTCGGGTGTATTTATGAC 1766
|||||
RESULT 8
ADN73150/C
ID ADN73150 standard; cDNA; 2931 BP.
XX
XX AC ADN73150;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 1045.
XX
XX KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
XX growth regulator; animal feed product; thale cress;
XX cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
XX OS Arabidopsis thaliana.
XX
XX PN WO2004035798-A2.
XX
XX PD 29-APR-2004.
XX
XX PF 20-OCT-2003; 2003WO-EP011658.
XX
XX PR 18-OCT-2002; 2002EP-00079408.
XX
XX PA (CROP-) CROPDESIGN NV.
XX
XX PI Inze D, De Veylder L, Vlieghe K;
XX
XX WIPI; 2004-348466/32.
XX
XX DR P-PSDB; ADN73151.
XX

```

PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.

XX Claim 1; SEQ ID NO 1045; 134pp; English.

XX This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/DPa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/DPa
CC transcription factor, given in an exemplification of the invention.

SQ Sequence 2931 BP; 818 A; 547 C; 682 G; 884 T; 0 U; 0 Other;

Query Match 65.3%; Score 19.6; DB 12; Length 2931;
Best Local Similarity 84.6%; Pred. No. 45;

Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 ATGTATCTCCGGTCTACATCTATGAC 30

Db 2591 AAGTACGTCTCGTGTACATCTATGAC 2566

RESULT 9

ACN44932_0
WP Sequence split into 4 fragments LOCUS ACN44932 Accession Acn44932

WP Fragment Name Begin End
WP ACN44932_0 1 110000
WP ACN44932_1 100001 210000
WP ACN44932_2 200001 310000
WP ACN44932_3 300001 374849
ID ACN44932 standard; DNA; 374849 BP.

XX ACN44932;

XX 18-NOV-2004 (first entry)

XX Mouse genomic sequence mCG5728.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX Mus musculus.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 1627; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC carcinoma. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

SQ Sequence 374849 BP; 106870A; 74478C; 74612G; 111149T; 0U; 77400Other;

Query Match 64.7%; Score 19.4; DB 11; Length 110000;

Best Local Similarity 79.3%; Pred. No. 1.1e+02;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGCGTGTACATCTATGA 29

Db 6156 AACAAATGATGTCGCGTGTACATCTATGA 6184

RESULT 10

ACN45116/c

ID ACN45116 standard; DNA; 42772 BP.

XX ACN45116;

XX 18-NOV-2004 (first entry)

XX Mouse genomic sequence mCG8527.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX Mus musculus.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 1903; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published
 XX

SQ Sequence 42772 BP; 11942 A; 8858 C; 9149 G; 11715 T; 0 U; 1108 Other;
 Query Match 64.0%; Score 19.2; DB 11; Length 42772;
 Best Local Similarity 87.5%; Pred. No. 1.2e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ATGTATGTCGGTGTCATCTATG 28

Db 15706 ATGTATGTCGGTGTCATCTATG 15683

RESULT 11

ACN51227
 ID ACN51227 standard; cDNA; 512 BP.

AC ACN51227;

DT 02-DEC-2004 (first entry)

DE Cotton androecium tissue EST Clone ID: LIB3828-012-Q1-N6-C3, SEQ:6008.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
 KW variety Nucotton33B; library LIB3828; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.

XX Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021123.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FING/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.

XX Claim 1; SEQ ID NO 6008; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucotton33B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining

CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes. Genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety Nucotton33B androecium tissue cDNA library (LIB3828). The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the US
 CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
 XX

SQ Sequence 512 BP; 159 A; 116 C; 91 G; 146 T; 0 U; 0 Other;

Query Match 63.3%; Score 19; DB 13; Length 512;

Best Local Similarity 81.5%; Pred. No. 63;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CAATGTATGTCGGTGTCATCTATGA 29

Db 394 CAATCCATTCTCGTGTCATCTCTGA 420

RESULT 12

AD853615/c

ID AD853615 standard; DNA; 2700 BP.

AC AD853615;

XX 04-DEC-2003 (first entry)

XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4157.

XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;

XX toxicity marker; toxicity progression; drug screening;

XX primary rat hepatocyte toxicity modelling; gene; ds.

OS Rattus norvegicus.

XX WO2003065993-A2.

XX 14-AUG-2003.

XX 04-FEB-2003; 2003WO-US003482.

XX 04-FEB-2002; 2002US-0353171P.

PR 13-MAR-2002; 2002US-0363534P.

PR 08-APR-2002; 2002US-0370248P.

PR 10-APR-2002; 2002US-0371134P.

PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.

PR 11-APR-2002; 2002US-0371413P.

PR 19-APR-2002; 2002US-0373601P.

PR 19-APR-2002; 2002US-0373602P.

PR 22-APR-2002; 2002US-0374139P.

PR 09-MAY-2002; 2002US-0378652P.

PR 09-MAY-2002; 2002US-0378653P.

PR 09-MAY-2002; 2002US-0378655P.

PR 09-JUL-2002; 2002US-0394230P.

PR 09-JUL-2002; 2002US-0394253P.

PR 04-SEP-2002; 2002US-0407688P.

PR 28-JAN-2003; 2003US-0442900P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;

PI Elashoff M;

XX WPI; 2003-731472/69.

KW antibody; gene therapy; neurodegenerative disease; eye disorder;
 KW cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;
 XX tooth abnormality; wound; ds.
 OS Homo sapiens.

Key Location/Qualifiers

FT exon 33104..33277
 FT /*tag= a
 FT /number= 2
 FT intron 33278..77746
 FT /*tag= b
 FT /number= 2
 FT exon 77747..77989
 FT /*tag= c
 FT /number= 3
 FT intron 77990..79103
 FT /*tag= d
 FT /number= 3
 FT exon 79104..79224
 FT /*tag= e
 FT /number= 4
 FT intron 79225..101022
 FT /*tag= f
 FT /number= 4
 FT exon 101023..101144
 FT /*tag= g
 FT /number= 5
 FT intron 101145..113377
 FT /*tag= h
 FT /number= 5
 FT exon 113378..113560
 FT /*tag= i
 FT /number= 6
 FT intron 113561..115985
 FT /*tag= j
 FT /number= 6
 FT exon 115986..116183
 FT /*tag= k
 FT /number= 7
 FT intron 116184..135707
 FT /*tag= l
 FT /number= 7
 FT exon 135708..135836
 FT /*tag= m
 FT /number= 8
 FT intron 135837..146471
 FT /*tag= m
 FT /number= 8
 FT exon 146472..146628
 FT /*tag= o
 FT /number= 9
 FT intron 146629..148761
 FT /*tag= p
 FT /number= 9
 FT exon 148762..148883
 FT /*tag= q
 FT /number= 10
 FT intron 148884..150044
 FT /*tag= r
 FT /number= 10
 FT exon 150045..150254
 FT /*tag= s
 FT /number= 11
 FT intron 150255..153815
 FT /*tag= t
 FT /number= 11
 FT exon 153816..154031
 FT /*tag= u
 FT /number= 12
 FT intron 154032..158580
 FT /*tag= v
 FT /number= 12

FT exon 158581..158802
 FT /*tag= w
 FT /number= 13
 FT intron 158803..173982
 FT /*tag= x
 FT /number= 13
 FT exon 173983..174177
 FT /*tag= y
 FT /number= 14
 FT intron 174178..181006
 FT /*tag= z
 FT /number= 14
 FT exon 181007..181129
 FT /*tag= aa
 FT /number= 15
 FT intron 181130..183612
 FT /*tag= ab
 FT /number= 15
 FT exon 183613..183800
 FT /*tag= ac
 FT /number= 16
 FT intron 183801..185152
 FT /*tag= ad
 FT /number= 16
 FT exon 185153..187765
 FT /*tag= ae
 FT /number= 17
 XX WO200138519-A1.
 XX 31-MAY-2001.
 XX 24-NOV-2000; 2000WO-AU001435.
 XX 26-NOV-1999; 99AU-00004348.
 XX (UYQU) UNIV QUEENSLAND.
 XX Little M, Yamada T, Holmes G, Georgas K, Kolle G, Wilkinson L;
 XX WPI; 2001-343951/36.
 XX Nucleic acids from human chromosome 2p21-16.3 and the encoded peptide,
 XX useful for preventing, diagnosing and treating e.g. eye disease,
 XX especially cataract formation.
 XX Claim 4; Fig 3; 169pp; English.
 XX The invention relates to nucleic acids from human chromosome 2p21-16.3
 XX and the encoded peptide (and mouse and chicken orthologues) that
 XX comprises a PGCCPLP group, an insulin-like growth factor binding protein
 XX (IGFBP)-like domain, cysteine-rich domains, an RGD (undefined) group and
 XX a transmembrane domain. The protein, e.g. CRIM1, interacts with peptides
 XX of the transforming growth factor superfamily. A composition comprising
 XX an expression construct comprising the nucleic acids of the invention or
 XX a mimetic which antagonises or mimics an activity of a CRIM1 polypeptide
 XX may be used in a method for modulating the biological activity of a
 XX polypeptide of the bone morphogenic protein (BMP) family. In this way
 XX they may be used to prevent or treat an eye disease, especially cataract
 XX formation. They may also be used to treat neurodegenerative diseases,
 XX renal and kidney disease, bone and tooth abnormalities, wounds and skin
 XX damage, e.g. by use of the nucleic acid in gene therapy by using
 XX antibodies directed against CRIM1 polypeptides. The present sequence is a
 XX Human genomic DNA containing exons 2-17 of the CRIM1 gene
 SQ Sequence 236303 BP; 70199 A; 46544 C; 47996 G; 71563 T; 0 U; 1 Other;
 Query Match 63.3%; Score 19; DB 4; Length 236303;
 Best Local Similarity 81.5%; Pred. No. 2.le+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 4 AATGTATGTCGGGTGATCATCTATGAC 30
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 133196 AATGTGTATTAGATGTACATCTGTGAC 133222

RESULT 16
AAF21611
ID AAF21611 standard; DNA; 349980 BP.
XX AC
XX AAF21611;
XX DT
XX 13-MAR-2001 (first entry)
XX DE
XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
XX KW
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX OS
XX Neisseria meningitidis.
XX WO200066791-A1.
XX FN
XX 09-NOV-2000.
XX PD
XX 08-MAR-2000; 2000WO-US005928.
XX PF
XX 30-APR-1999; 99US-0132068P.
XX PR
XX 08-OCT-1999; 99WO-US023573.
XX PR
XX 28-FEB-2000; 2000GB-00004695.
XX PA
XX (CHIR) CHIRON CORP.
XX PA (GENO-) INST GENOMIC RES.
XX XX
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Frazer CM, Grandi G;
XX WPI; 2000-647603/62.
XX DR
XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.
PT
PS Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used
XX SQ
XX Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;
Query Match 63.3%; Score 19; DB 3; Length 349980;
Best Local Similarity 81.5%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 1 AACAAATGTATGTCGGGTGATCATCTAT 27
|||||

Db 52042 AACAAATGTATTTCTATTGTTCTATCATCTAT 52068

RESULT 17
AAS62657/C
ID AAS62657 standard; cDNA; 1630 BP.
XX AC
XX AAS62657;
XX DT
XX 14-FEB-2002 (first entry)
XX DE
XX cDNA sequence #444 encoding novel human secreted protein.
XX KW
XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX OS
XX Homo sapiens.
XX WO200177291-A2.
XX FN
XX 18-OCT-2001.
XX PD
XX 29-MAR-2001; 2001WO-US010485.
XX PF
XX 06-APR-2000; 2000US-0195604P.
XX PR
XX (GEMY) GENETICS INST INC.
XX PA
XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
PI WPI; 2002-010900/01.
XX DR
XX New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease.
PT
PS Claim 1; Page 308; 391pp; English.
XX
XX The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The
CC proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are useful
CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
CC immune deficiency disorders (e.g. severe combined immunodeficiency
CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis) and
CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
CC the invention are also useful in gene therapy. AAS62214-AAS62838
CC represent the cDNA sequences of the invention that encode for novel human
CC secreted proteins
XX SQ
XX Sequence 1630 BP; 409 A; 346 C; 284 G; 590 T; 0 U; 1 Other;
Query Match 62.7%; Score 18.8; DB 6; Length 1630;
Best Local Similarity 76.7%; Pred. No. 98;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Oy 1 AACAAATGTATGTCGGGTGATCATCTATGAC 30
|||||

Db 633 AACGAAGTAAGCCCTGTGTATCATCTATGAC 604

RESULT 18
AAA30829
ID AAA30829 standard; DNA; 1569 BP.
XX AC
XX AAA30829;
AC

```

XX 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
XX Zebrafish PTH3R receptor coding sequence.
XX
XX Zebrafish; PTH1R receptor; PTH3R receptor; diagnosis; cancer;
KW parathyroid hormone type 1 receptor; parathyroid hormone type 3 receptor;
KW ds.
XX
XX Danio rerio.
XX
XX Key Location/Qualifiers
XX CDS l..1569
XX /*tag= a
XX /product= "PTH3R receptor"
XX /partial
XX /note= "no stop codon given"
XX
XX WO200032771-A1.
XX
XX 08-JUN-2000.
XX
XX 28-MAY-1999; 99WO-US011883.
XX
XX 30-NOV-1998; 98US-0110467P.
XX
XX (JUEP/) JUEPPNER H.
XX (RUBI/) RUBIN D A.
XX
XX Jueppner H, Rubin DA;
XX
XX WPI; 2000-412319/35.
XX P-PSDE; AAY90231.
XX
XX Novel zebrafish parathyroid hormone/parathyroid hormone related peptide
XX receptor 3 and isolated nucleic acid encoding zebrafish parathyroid
XX hormone receptor 1 for treating disorders associated with receptor
XX function.
XX
XX Claim 23; Fig 1d; 11lpp; English.
XX
XX This sequence encodes a parathyroid hormone receptor type 3 (PTH3R)
XX receptor protein of the invention. The invention also relates to a PTH1R
XX receptor protein. Antagonists of PTH1R or PTH3R can be used for the
XX treatment of diseases associated with an increase in PTH1R or PTH3R
XX activity, respectively. The peptides are used for diagnosis or prognosis
XX of diseases and disorders associated with PTH3R or PTH1R, such as cancer.
XX The polypeptides can be used as a molecular weight markers on sodium
XX dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels, or
XX on molecular sieve gel filtration columns. Antigenic epitope-bearing
XX peptides and polypeptides are useful to raise antibodies, including
XX monoclonal antibodies, that bind specifically to a polypeptide. The
XX peptides are useful during diagnosis of diseases and disorders in mammals
XX involving PTH1R or PTH3R receptor expression or function. Mutations that
XX affect PTH1R or PTH3R sequence and/or expression levels of PTH1R or PTH3R
XX could be diagnostic for patients with disease or disorders of a
XX developmental, physiological or neurological nature. The nucleic acid
XX molecules are valuable for chromosome identification. The mapping of DNAs
XX to chromosomes is an important first step in correlating those sequences
XX with genes associated with disease. (Updated on 15-SEP-2003 to
XX standardise OS field)
XX
XX Query Match 62.0%; Score 18.6; DB 3; Length 1569;
XX Best Local Similarity 84.0%; Pred. No. 1.2e+02;
XX Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 6 TGTATGTCGGGTGTACATCTATGAC 30
XX 246 TCTGTGCTCTGAGTACATCTATGAC 270
XX
XX
XX RESULT 20
XX ADC42329
XX ID ADC42329 standard; cDNA; 1671 BP.
XX
XX ADC42329;
XX
XX 18-DEC-2003 (first entry)
XX

```

```

XX
XX RESULT 19
XX AAA48445
XX ID AAA48445 standard; cDNA; 1671 BP.
XX
XX AAA48445;
XX
XX 15-SEP-2003 (revised)
DT 27-OCT-2000 (first entry)
XX
XX Zebrafish parathyroid hormone type-3 receptor PTH3R coding sequence.
XX
XX Zebrafish; parathyroid hormone type-3 receptor; PTH3R;
KW developmental disorder; physiological disorder; neurological disorder;
KW ss.
XX
XX Danio rerio.
XX
XX Key Location/Qualifiers
XX CDS 97..1671
XX /*tag= a
XX /product= "PTH3R"
XX /partial
XX
XX WO200032775-A1.
XX
XX 08-JUN-2000.
XX
XX 30-NOV-1999; 99WO-US028207.
XX
XX 30-NOV-1998; 98US-0110467P.
XX
XX (JUEP/) JUEPPNER H.
XX (RUBI/) RUBIN D A.
XX
XX Jueppner H, Rubin DA;
XX
XX WPI; 2000-412323/35.
XX
XX New nucleic acids encoding parathyroid hormone receptors PTH1R and PTH3R,
XX useful for treating diseases or disorders associated with impaired
XX receptor functions comprises a specific nucleotide sequence.
XX
XX Disclosure; Fig 1C; 11lpp; English.
XX
XX The present sequence is the parathyroid hormone type-3 receptor (PTH3R)
XX coding sequence from the zebrafish. It was obtained by sequencing a cDNA
XX clone. The gene and protein can be used to detect diseases in man where
XX the receptor is either overexpressed or underexpressed, and they can be
XX used to treat these diseases, which may be developmental, physiological
XX or neurological disorders. They can also be used to identify agonists and
XX antagonists which can be used in a similar manner. In addition, the gene
XX can be used for chromosome identification. (Updated on 15-SEP-2003 to
XX standardise OS field)
XX
XX Query Match 62.0%; Score 18.6; DB 3; Length 1671;
XX Best Local Similarity 84.0%; Pred. No. 1.2e+02;
XX Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 6 TGTATGTCGGGTGTACATCTATGAC 30
XX 342 TCTGTGCTCTGAGTACATCTATGAC 366
XX
XX
XX RESULT 20
XX ADC42329
XX ID ADC42329 standard; cDNA; 1671 BP.
XX
XX ADC42329;
XX
XX 18-DEC-2003 (first entry)
XX

```

```
XX cDNA encoding zebrafish parathyroid hormone receptor PTH3R seq id 26.
XX parathyroid hormone; PTH; PTH-related peptide; PTHrP;
XX parathyroid hormone receptor; PTHR; chromosome identification; zebrafish;
XX PTHrR, ss.
XX
XX Danio rerio.
XX
XX Key Location/Qualifiers
XX CDS 97..1668
XX /*tag= a
XX /product= "zebrafish PTH3R"
XX /note= "parathyroid hormone receptor"
XX
XX US6541220-B1.
XX
XX 01-APR-2003.
XX
XX 30-NOV-1999; 99US-00449632.
XX
XX 30-NOV-1998; 98US-0110467P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Jueppner H, Rubin DA;
XX
XX WPI; 2003-754511/71.
XX P-PSDB; ADC42307.
XX
XX Novel nucleic acid comprising a polynucleotide encoding parathyroid
XX hormone/parathyroid hormone-related peptide receptor, useful for
XX chromosome identification.
XX
XX Disclosure; SEQ ID NO 26; 52pp; English.
XX
XX The invention describes an isolated nucleic acid (I) comprising a
XX polynucleotide having a nucleotide sequence chosen from nucleotide
XX sequence encoding a parathyroid hormone (PTH)/PTH-related peptide (PTHrP)
XX receptor (PTHrR receptor) having a fully defined sequence of 536 amino
XX acids as given in the specification, PTHrR receptor, mature PTHr
XX receptor, PTHrR extracellular or transmembrane domain, and their
XX complement. (I) is useful for diagnosing and treating decrease in the
XX standard or normal level of PTHrR receptor activity in an individual, and
XX for chromosome identification. This sequence encodes zebrafish PTH3R.
XX
XX Sequence 1671 BP; 409 A; 381 C; 449 G; 432 T; 0 U; 0 Other;
XX
XX Query Match 62.0%; Score 18.6; DB 10; Length 1671;
XX Best Local Similarity 84.0%; Pred. No. 1.2e+02;
XX Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 6 TGTATGTCGGGTGACATCTATGAC 30
XX 342 TCTGTGTCCTGAGTACATCTATGAC 366
XX
XX RESULT 21
XX AAA48446
XX ID AAA48446 standard; cDNA; 2152 BP.
XX
XX AAA48446;
XX
XX 15-SEP-2003 (revised)
XX 27-OCT-2000 (first entry)
XX
XX Zebrafish parathyroid hormone type-3 receptor PTH3R coding sequence.
XX
XX DE
XX Zebrafish; parathyroid hormone type-3 receptor; PTH3R;
XX developmental disorder; physiological disorder; neurological disorder;
XX ss.
XX
XX Danio rerio.
XX
```

```
XX Key Location/Qualifiers
XX CDS 394..2022
XX /*tag= a
XX /product= "PTH3R"
XX
XX WO200032775-A1.
XX
XX 08-JUN-2000.
XX
XX 30-NOV-1999; 99WO-US028207.
XX
XX 30-NOV-1998; 98US-0110467P.
XX
XX (JUEP/) JUEPPNER H.
XX (RUBI/) RUBIN D A.
XX
XX Jueppner H, Rubin DA;
XX
XX WPI; 2000-412323/35.
XX P-PSDB; AAY99601.
XX
XX New nucleic acids encoding parathyroid hormone receptors PTHrR and PTH3R,
XX useful for treating diseases or disorders associated with impaired
XX receptor functions comprises a specific nucleotide sequence.
XX
XX Claim 23; Fig 1D; 11pp; English.
XX
XX The present sequence is the parathyroid hormone type-3 receptor (PTH3R)
XX coding sequence from the zebrafish. It was obtained by sequencing a cDNA
XX clone. The gene and protein can be used to detect diseases in man where
XX the receptor is either overexpressed or underexpressed, and they can be
XX used to treat these diseases, which may be developmental, physiological
XX or neurological disorders. They can also be used to identify agonists and
XX antagonists which can be used in a similar manner. In addition, the gene
XX can be used for chromosome identification. (Updated on 15-SEP-2003 to
XX standardise OS field)
XX
XX Sequence 2152 BP; 592 A; 471 C; 529 G; 559 T; 0 U; 1 Other;
XX
XX Query Match 62.0%; Score 18.6; DB 3; Length 2152;
XX Best Local Similarity 84.0%; Pred. No. 1.3e+02;
XX Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 6 TGTATGTCGGGTGACATCTATGAC 30
XX 639 TCTGTGTCCTGAGTACATCTATGAC 663
XX
XX RESULT 22
XX ADC42306
XX ID ADC42306 standard; cDNA; 2152 BP.
XX
XX ADC42306;
XX
XX 18-DEC-2003 (first entry)
XX
XX cDNA encoding zebrafish parathyroid hormone receptor PTH3R seq id 3.
XX
XX parathyroid hormone; PTH; PTH-related peptide; PTHrP;
XX parathyroid hormone receptor; PTHR; chromosome identification; zebrafish;
XX PTHrR; gene; ss.
XX
XX Danio rerio.
XX
XX Key Location/Qualifiers
XX CDS 394..2022
XX /*tag= a
XX /product= "zebrafish PTH3R"
XX /note= "Parathyroid hormone receptor"
XX
XX US6541220-B1.
XX
```

PD 01-APR-2003.
 XX
 PF 30-NOV-1999; 99US-00449632.
 XX
 PR 30-NOV-1998; 98US-0110467P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 XX
 XX Jueppner H, Rubin DA;
 XX
 XX WPI; 2003-754511/71.
 DR P-PSDB; ADC42307.
 XX
 XX Novel nucleic acid comprising a polynucleotide encoding parathyroid
 PT hormone/parathyroid hormone-related peptide receptor, useful for
 PT chromosome identification.
 XX
 XX Disclosure; SEQ ID NO 3; 52pp; English.
 XX
 CC The invention describes an isolated nucleic acid (I) comprising a
 CC polynucleotide having a nucleotide sequence chosen from nucleotide
 CC sequence encoding a parathyroid hormone (PTH)/PTH-related peptide (PTHrP)
 CC receptor (PTHrR) having a fully defined sequence of 536 amino
 CC acids as given in the specification, PTHrR receptor, mature PTHrR
 CC receptor, PTHrR extracellular or transmembrane domain, and their
 CC complement. (I) is useful for diagnosing and treating decrease in the
 CC standard or normal level of PTHrR receptor activity in an individual, and
 CC for chromosome identification. This sequence encodes zebrafish PTH3R.
 XX
 SQ Sequence 2152 BP; 592 A; 470 C; 529 G; 560 T; 0 U; 1 Other;
 Query Match 62.0%; Score 18.6; DB 10; Length 2152;
 Best Local Similarity 84.0%; Pred. No. 1.3e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 6 TGTATGTCGGGTGACATCTATGAC 30
 DB 639 TCTGTGTCCTGAGTACATCTATGAC 663
 RESULT 23
 ADH61060
 ID ADH61060 standard; cDNA; 2152 BP.
 XX
 AC ADH61060;
 XX
 DT 25-MAR-2004 (first entry)
 XX
 DE Zebrafish PTH3R cDNA from clone zPTH3R.
 XX
 KW Zebrafish; parathyroid hormone; PTH; parathyroid hormone related protein;
 KW PTHrP; parathyroid hormone related protein receptor; PTHrR; PTH3R;
 KW diagnosis; prognosis; pharmaceutical composition; chromosome assay; gene;
 KW ss.
 XX
 OS Danio rerio.
 XX
 FH Key Location/Qualifiers
 FT CDS 394..2022
 FT /*tag= b
 FT /product= "Zebrafish PTH3R protein"
 FT sig_peptide 394..457
 FT /*tag= a
 FT mat_peptide 458..2019
 FT /*tag= c
 FT /product= "Zebrafish mature PTH3R protein"
 XX
 DN US2003162256-A1.
 XX
 PD 28-AUG-2003.
 XX
 XX 25-FEB-2003; 2003US-00372095.
 PF
 XX

PR 30-NOV-1998; 98US-0110467P.
 PR 30-NOV-1999; 99US-00449632.
 XX
 PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
 XX
 XX Jueppner H, Rubin DA;
 XX
 XX WPI; 2003-897927/82.
 DR P-PSDB; ADH61061.
 DR
 XX New parathyroid hormone receptors designated PTHrR and PTH3R isolated
 PT from zebrafish are useful to diagnose and treat parathyroid hormone
 PT receptor-related diseases.
 XX
 XX Claim 23; SEQ ID NO 3; 53pp; English.
 XX
 CC The present invention relates to novel parathyroid hormone (PTH) and
 CC parathyroid hormone related protein (PTHrP) receptors PTHrR and PTH3R
 CC isolated from zebrafish. The invention is useful in the diagnosis and
 CC prognosis of certain diseases and disorders that express significantly
 CC decreased levels of PTHrR and PTH3R. The invention is also useful in
 CC preparing pharmaceutical compositions and in chromosome assays. The
 CC present sequence is zebrafish PTH3R cDNA.
 XX
 SQ Sequence 2152 BP; 592 A; 470 C; 529 G; 560 T; 0 U; 1 Other;
 Query Match 62.0%; Score 18.6; DB 10; Length 2152;
 Best Local Similarity 84.0%; Pred. No. 1.3e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 6 TGTATGTCGGGTGACATCTATGAC 30
 DB 639 TCTGTGTCCTGAGTACATCTATGAC 663
 RESULT 24
 ADA03077
 ID ADA03077 standard; DNA; 90442 BP.
 XX
 AC ADA03077;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Mouse mCG2257 carcinoma associated gene, SEQ ID NO:1595.
 XX
 KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ds.
 XX
 OS Mus sp.
 XX
 PN WO2003057146-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 26-DEC-2002; 2002WO-US041414.
 XX
 PR 26-DEC-2001; 2001US-00035832.
 XX
 PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW;
 XX
 XX WPI; 2003-587068/55.
 DR
 XX New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 PT
 XX Claim 1; SEQ ID NO 1595; 245pp; English.
 PS
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC

CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed murine CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 90442 BP; 21913 A; 21998 C; 23663 G; 22868 T; 0 U; 0 Other;

Query Match 62.0%; Score 18.6; DB 9; Length 90442;

Best Local Similarity 84.0%; Pred. No. 2.7e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 ATGTATGTCCGGTGACATCTATGA 29

||||||| ||| ||| ||| ||| ||| |||

Db 3731 ATGTATGTCCGGGCACATATATGA 3755

RESULT 25

ADA66361

ID ADA66361 standard; DNA; 90442 BP.

XX AC ADA66361;

XX DT 20-NOV-2003 (first entry)

XX DE Mouse mCG2257 gene genomic DNA sequence.

XX carcinoma-associated gene; CA gene; Rorc gene; mCG15938 gene; BAT1 gene;
 KW Iggap1 gene; IQGAP1 gene; Zfp29 gene; hCG27579 gene; Kcnj9 gene;
 KW Kcnj9 gene; Ppp3cc gene; Ppp3cc gene; mCG9110 gene; hCG27579 gene;
 KW cancer cell; lymphatic cell; breast cell; prostate cell; epithelial cell;
 KW carcinoma-associated protein; CAP; cytostatic; gene therapy; anticancer;
 KW vaccine; carcinoma; lymphoma carcinoma; lymphatic cancer; breast cancer;
 KW prostate cancer; DNA vaccine; animal model; mouse; murine; ds; mCG2257.

XX OS Mus sp.

XX PN W02003053224-A2.

XX PD 03-JUL-2003.

XX PF 20-DEC-2002; 2002WO-US041776.

XX PR 20-DEC-2001; 2001US-00034650.

XX PA (SAGR-) SAGRES DISCOVERY,

XX PI Morris DW, Engelhard EK;

XX DR WPI; 2003-569168/53.

XX Novel recombinant carcinoma-associated nucleic acid, useful for
 PT evaluating the effect of a candidate carcinoma drug, and for diagnosing
 PT carcinoma.

XX PS Claim 1; Page 161-174; 229pp; English.

XX This invention relates to a novel recombinant carcinoma-associated (CA)
 CC nucleic acid comprising a fully defined genomic, mRNA or coding sequences

CC of mouse Rorc gene or human RORC gene, mouse mCG15938 or human gene BAT1,
 CC mouse Iggap1 gene or human IQGAP1 gene, mouse Zfp29 gene or human
 CC hCG27579 gene, mouse Kcnj9 gene or human Kcnj9 gene, mouse Ppp3cc gene or
 CC human Ppp3cc gene, mouse mCG9110 gene or human hCG27579 gene, as given in
 CC the specification. CA genes are genes which are preferably expressed in
 CC cancer cells, preferably lymphatic, breast, prostate or epithelial cells.
 CC A compound which modifies the expression of the CA genes or bind to
 CC carcinoma-associated proteins (CAP) may have cytostatic activity and the
 CC sequences of the invention may enable the use of gene therapy or a
 CC development of an anticancer vaccine. Therefore the invention may be
 CC useful for diagnosis and treatment of carcinomas, especially lymphoma
 CC carcinoma, breast cancer and prostate cancer. The CA genes may also be
 CC useful as DNA vaccines and for generating animal models of carcinomas.
 CC The present sequence is that of the mouse mCG2257 gene genomic DNA
 CC sequence of the invention.

XX SQ Sequence 90442 BP; 21913 A; 21998 C; 23663 G; 22868 T; 0 U; 0 Other;

Query Match 62.0%; Score 18.6; DB 9; Length 90442;

Best Local Similarity 84.0%; Pred. No. 2.7e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 ATGTATGTCCGGTGACATCTATGA 29

||||||| ||| ||| ||| ||| ||| |||

Db 3731 ATGTATGTCCGGGCACATATATGA 3755

Search completed: October 28, 2005, 10:26:32

Job time : 246 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 07:48:14 ; Search time 1537 Seconds
(without alignments)
742.959 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgtatgtccggtgtacatcatgac 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23.2	77.3	763	8	B11833
C 2	21.2	70.7	416	8	A2227906
C 3	21.2	70.7	773	7	CO958388
C 4	21.2	70.7	815	9	AG498011
C 5	20.8	69.3	631	8	BZ666678
C 6	20.8	69.3	977	9	CL989513
C 7	20.6	68.7	631	5	BX507610
C 8	20.6	68.7	672	5	BX507786
C 9	20.4	68.0	643	6	CA501775
C 10	20.4	68.0	707	4	B3301068
C 11	20.4	68.0	862	9	AG347593
C 12	20.4	68.0	960	9	CL290684
C 13	20.2	66.7	390	8	BH204000
C 14	20.2	66.7	538	9	CL412828
C 15	20.2	66.7	542	8	AQ879634
C 16	20.2	66.7	605	8	AQ525705
C 17	20.2	66.7	773	8	AQ856387
C 18	20.2	66.7	858	7	CV480979
C 19	20.2	66.7	916	8	AZ204484
C 20	20.2	66.7	1330	8	B12483
C 21	19.6	65.3	266	8	BZ091289
C 22	19.6	65.3	380	1	AA751822
C 23	19.6	65.3	385	8	BH121074
C 24	19.6	65.3	396	7	CF889471

25	19.6	65.3	475	1	AI562331
C 26	19.6	65.3	561	8	BZ251753
C 27	19.6	65.3	658	8	CC181465
C 28	19.6	65.3	700	9	CL716190
C 29	19.6	65.3	905	9	CL507353
C 30	19.4	64.7	360	2	BF249864
C 31	19.4	64.7	365	5	BW055193
C 32	19.4	64.7	400	7	CV166681
C 33	19.4	64.7	415	8	AQ205261
C 34	19.4	64.7	430	7	CV166144
C 35	19.4	64.7	540	8	BH180506
C 36	19.4	64.7	540	9	CNS07MCG
C 37	19.4	64.7	568	7	CV167277
C 38	19.4	64.7	606	8	AZ821971
C 39	19.4	64.7	632	5	BU411840
C 40	19.4	64.7	634	6	CA256318
C 41	19.4	64.7	670	8	BZ366040
C 42	19.4	64.7	678	8	BZ367686
C 43	19.4	64.7	706	8	AZ449367
C 44	19.4	64.7	741	8	BZ366005
C 45	19.4	64.7	815	8	BZ720063
C 46	19.4	64.7	827	8	BZ119859
C 47	19.4	64.7	845	9	CG938893
C 48	19.4	64.7	874	4	BG621981
C 49	19.4	64.7	877	2	BF179970
C 50	19.4	64.7	1047	9	CL991786
C 51	19.4	64.7	1360	9	AG321305
C 52	19.2	64.0	677	8	AZ597359
C 53	19.2	64.0	681	8	AZ983790
C 54	19.2	64.0	698	9	CR100026
C 55	19.2	64.0	715	5	CG106613
C 56	19.2	64.0	747	9	BX750750
C 57	19.2	64.0	809	8	CC443650
C 58	19.2	64.0	810	5	BX755052
C 59	19.2	64.0	814	8	BZ962341
C 60	19.2	64.0	820	5	BX761914
C 61	19.2	64.0	822	5	BX747963
C 62	19.2	64.0	825	9	CG093988
C 63	19.2	64.0	863	5	BX781945
C 64	19.2	64.0	911	9	CC976050
C 65	19.2	64.0	916	8	CC431604
C 66	19.2	64.0	924	5	BX782389
C 67	19.2	64.0	932	9	CG943126
C 68	19.2	64.0	943	9	CG204329
C 69	19.2	64.0	1080	9	CC807393
C 70	19.2	64.0	1345	9	AG294534
C 71	19.2	63.3	289	9	AG213543
C 72	19.2	63.3	311	8	AZ130373
C 73	19.2	63.3	373	6	BY650351
C 74	19.2	63.3	432	1	AL924646
C 75	19.2	63.3	475	2	AW594793
C 76	19.2	63.3	501	8	AZ733175
C 77	19.2	63.3	510	6	CB782681
C 78	19.2	63.3	510	8	AZ152910
C 79	19.2	63.3	513	9	CNS04LGK
C 80	19.2	63.3	515	8	AZ180742
C 81	19.2	63.3	537	9	CC639179
C 82	19.2	63.3	540	7	CN603311
C 83	19.2	63.3	556	7	CN603555
C 84	19.2	63.3	556	7	CN603648
C 85	19.2	63.3	571	7	CN603208
C 86	19.2	63.3	578	7	CN603175
C 87	19.2	63.3	597	7	CN602488
C 88	19.2	63.3	597	7	CN602534
C 89	19.2	63.3	597	7	CN603063
C 90	19.2	63.3	603	8	AZ980570
C 91	19.2	63.3	609	7	CN602628
C 92	19.2	63.3	614	7	CN602549
C 93	19.2	63.3	615	8	BZ984614
C 94	19.2	63.3	638	9	BX221462
C 95	19.2	63.3	647	6	CD777639
C 96	19.2	63.3	651	8	BM485271
C 97	19.2	63.3	693	8	BZ624616

AI562331	TENS2326
BZ251753	CH230-306
CC181465	JHpa_1_A2
CL716190	OR_BB004
CL507353	SAL1780
BF249864	pa87508_Y
BW055193	BM055193
CV166681	rmsexl_00
AQ205261	HS_3221_A
CV166144	rmsexl_00
BH180506	016_P_04
AL617458	T3_end_of
CV167277	rmsexl_00
AZ821971	2M0094E19
BU411840	602954387
CA256318	SCSGFLA18
BZ366040	ic94c01.9
BZ367686	ld06a02.g
AZ449367	1M0247N05
BZ366005	ic94a01.9
BZ720063	PUCFO79TD
BZ119859	CH230-422
CG938893	MBEHL56TF
BG621981	602646561
BF179970	601806315
CL991786	ZMMBHf000
AG321305	Mus muscu
AZ597359	1M0411K20
AZ983790	2M0265SM05
CR100026	Forward s
BX750750	BX750750
CG106613	PURAR70TD
CC443650	PUHNS72TB
BX755052	BX755052
BZ962341	PUCAJ75FD
BX761914	BX761914
BX747963	BX747963
CG093988	PUJFG74TD
BX781945	BX781945
CC976050	ZUAR13TV
CC431604	FUHLR90TB
BX782389	BX782389
CG943126	MBEHE31TF
CG204329	PUJEB68TD
CC807393	ZMMBCC051
AG294534	Mus muscu
AG213543	Oryza sat
AZ130373	OSJNB010
BY650351	BY650351
AL924646	AL924646
AW594793	fk24b06_Y
AZ733175	RPCI-24-1
CB782681	AMGNNUC:N
AZ152910	SP_0028_A
CNS04LGK	Tetraodon
AZ180742	SP_0181_A
CC639179	OGHWM70TV
CN603311	C0006786N
CN603555	C0006787K
CN603648	C0006787P
CN603208	C0006786I
CN603175	C0006786G
CN602488	C0006784B
CN602534	C0006784E
CN603063	C0006786A
AZ980570	2M0257108
CN602628	C0006784J
CN602549	C0006784F
BZ984614	PURAY10TD
BX221462	Danio rer
CD777639	TDSUBS_RZ
BM485271	pgmln.pk0
BZ624616	ig92c12.b

98 19 63.3 709 4 BG545083 BG545083 602572271
 99 19 63.3 723 5 BU459209 BU459209 603367031
 100 19 63.3 743 7 CV113883 CV113883 AGENCOURT

ALIGNMENTS

RESULT 1
 B11833/c
 LOCUS
 DEFINITION B11833 F15H3-T7 IGF Arabidopsis thaliana genomic clone F15H3, genomic survey sequence.
 B11833
 VERSION B11833.1 GI:2092953
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 763)
 Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.

REFERENCE
 AUTHORS BAC End Sequences at ATGC
 TITLE Unpublished (1997)
 JOURNAL Other GSSs: F15H3-Sp6.1, F15H3-T7.1, F15H3-Sp6
 COMMENT Contact: Ecker J.
 Arabidopsis Thaliana Genome Center
 University of Pennsylvania
 Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104

TE1: 215-898-9384
 FAX: 215-898-8780
 Email: jecker@atgenome.bio.upenn.edu
 Seq primer: T7

Class: BAC ends
 High quality sequence start: 322
 High quality sequence stop: 343.
 Location/Qualifiers

1..763
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="F15H3"
 /sex="hermaphrodite"
 /clone_lib="IGF"
 /note="Vector: BoloBACII; Site_1: EcoRI; Site_2: EcoRI;
 Produced by Thomas Altmann"

Query Match 77.3%; Score 23.2; DB 8; Length 763;
 Best Local Similarity 89.3%; Pred. No. 8.7;
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATGA 29
 |||||
 Db 130 ACGATGATGTCCTGTGTACACCTATGA 103

RESULT 2
 A2227906
 LOCUS
 DEFINITION A2227906 RPCI-23-82A18.TJ RPCI-23 Mus musculus genomic clone RPCI-23-82A18, genomic survey sequence.
 A2227906
 VERSION A2227906.1 GI:8535955
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS BAC End Sequences at ATGC
 TITLE Unpublished (1997)
 JOURNAL Other GSSs: F15H3-Sp6.1, F15H3-T7.1, F15H3-Sp6
 COMMENT Contact: Ecker J.
 Arabidopsis Thaliana Genome Center
 University of Pennsylvania
 Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104

TE1: 215-898-9384
 FAX: 215-898-8780
 Email: jecker@atgenome.bio.upenn.edu
 Seq primer: T7

Class: BAC ends
 High quality sequence start: 322
 High quality sequence stop: 343.
 Location/Qualifiers

1..763
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="F15H3"
 /sex="hermaphrodite"
 /clone_lib="IGF"
 /note="Vector: BoloBACII; Site_1: EcoRI; Site_2: EcoRI;
 Produced by Thomas Altmann"

Query Match 77.3%; Score 23.2; DB 8; Length 763;
 Best Local Similarity 89.3%; Pred. No. 8.7;
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATGA 29
 |||||
 Db 130 ACGATGATGTCCTGTGTACACCTATGA 103

RESULT 3
 CO958388/c
 LOCUS
 DEFINITION CO958388 AGENCOURT_30676684 NIH_ZGC_14 Danio rerio cDNA clone IMAGE:7405493 5', mRNA sequence.

CO958388
 VERSION CO958388.1 GI:51322970
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov

Query Match 70.7%; Score 21.2; DB 8; Length 416;
 Best Local Similarity 88.5%; Pred. No. 67;
 Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CAATGATGTCGGGTGACATCTATG 28
 |||||
 Db 204 CTATGATGTCGGGTGACATCTATG 229

RESULT 3
 CO958388/c
 LOCUS
 DEFINITION CO958388 AGENCOURT_30676684 NIH_ZGC_14 Danio rerio cDNA clone IMAGE:7405493 5', mRNA sequence.

CO958388
 VERSION CO958388.1 GI:51322970
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov

Query Match 70.7%; Score 21.2; DB 8; Length 416;
 Best Local Similarity 88.5%; Pred. No. 67;
 Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CAATGATGTCGGGTGACATCTATG 28
 |||||
 Db 204 CTATGATGTCGGGTGACATCTATG 229

RESULT 3
 CO958388/c
 LOCUS
 DEFINITION CO958388 AGENCOURT_30676684 NIH_ZGC_14 Danio rerio cDNA clone IMAGE:7405493 5', mRNA sequence.

CO958388
 VERSION CO958388.1 GI:51322970
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov

REFERENCE

1 (bases 1 to 416)
 Zhao, S., Nierman, M., Feldblyum, T., Malek, J., Shatsman, S., Akınret, B., Levine, M., McGann, S., Tsengaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT


```

/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match      68.7%; Score 20.6; DB 5; Length 672;
Best Local Similarity 85.2%; Pred. No. 1.4e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  2  ACAATGATATCCGGTGATACATCTATG 28
    ||| ||||| ||||| ||||| |||||
Db 156 AAAACGATATCCAGTGATACATTTATG 130

RESULT 9
CA501775/c
LOCUS
DEFINITION
WHE4038_E11_I222T Wheat meiotic anther cDNA library Triticum
aestivum cDNA clone WHE4038_E11_I22, mRNA sequence.
CA501775
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 643)
REFERENCE
AUTHORS
Anderson,O.D., Chao,S., Crossman,C., Langridge,P., Lazo,G.R.,
Pham,J., Rausch,C.J., Sutton,T., Woo,J. and Wilson,C.
TITLE
The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
JOURNAL
Unpublished (2002)
COMMENT
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: T7 primer.
FEATURES
source
Location/Qualifiers
1..643
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4038_E11_I22"
/tissue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants
were grown in a glasshouse. Anther meiotic stage was
determined by removing anthers from individual primary
florets. One anther was sacrificed for microscopic
staging, and if determined to be between (and including)
meiotic stages pre-meiosis and metaphase I, the remaining
two anthers were collected and pooled for library
construction. The tissue, total RNA, and poly(A) RNA were
prepared, cDNA synthesised, and directionally ligated into
pSPORT1 by Tim Sutton in the P Langridge Lab at the
Department of Plant Science, University of Adelaide, Waite
Campus, Australia. Average insert size 1.5kb. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
```

Query Match

68.0%; Score 20.4; DB 6; Length 643;

Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACATGATATCCGGTGATACATCTATGAC 30
||| ||||| ||||| ||||| |||||
Db 592 AACTTTGCATGTCGGGTTAACAATGAC 563

RESULT 10
BJ301068/c
LOCUS
DEFINITION
BJ301068 Y. Ogihara unpublished cDNA library, wh_yd Triticum
aestivum cDNA clone whyd5n12 5', mRNA sequence.
BJ301068
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 707)
REFERENCE
AUTHORS
Ogihara,Y. and Murai,K.
TITLE
Expressed genes in Triticum aestivum
JOURNAL
Unpublished (2002)
COMMENT
Contact: Tadasu Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
Location/Qualifiers
1..707
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whyd5n12"
/tissue_type="spikelet at late flowering"
/dev_stage="Freeke's scale 6"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_yd"

Query Match 68.0%; Score 20.4; DB 4; Length 707;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACATGATATCCGGTGATACATCTATGAC 30
||| ||||| ||||| ||||| |||||
Db 222 AACTTTGCATGTCGGGTTAACAATGAC 193

RESULT 11
AG347593/c
LOCUS
DEFINITION
AG347593
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Mus musculus molossinus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
REFERENCE
AUTHORS
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE
BAC end Sequences of Library MSMg01
JOURNAL
Unpublished
REFERENCE
AUTHORS
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE
Direct Submision
JOURNAL
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical


```

Qy 3 CAATGTATGTCGGGTGACATCTATGAC 30
    ||||| ||||| ||||| ||||| |||||
Db 387 CAATGTTGTCGGGTGATGATGAC 360

RESULT 14
CL412828
LOCUS
DEFINITION
  RPCI44_431D15.r RPCI-44 Sus scrofa genomic clone RPCI44_431D15,
  genomic survey sequence.
ACCESSION
  CL412828
VERSION
  CL412828.1 GI:51446888
SOURCE
  GSS.
ORGANISM
  Sus scrofa (pig)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
  1 (bases 1 to 538)
  Rogatcheva, M.B., Meyers, S., He, W., Larkin, D.M., Marron, B.M.,
  Beaver, J.E. and Schook, L.B.
  Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
  Through Comparative Genomics
  Unpublished (2004)
JOURNAL
  Other GSSs: RPCI44_431D15.f
COMMENT
  Contact: Lawrence B. Schook
  Department of Animal Sciences
  University of Illinois at Urbana Champaign
  1201 W. Gregory Dr., Urbana, IL 61801, USA
  Tel: 217 265 5326
  Fax: 217 244 5617
  Email: schook@uiuc.edu
  Clones are derived from the porcine BAC library RPCI-44
  (http://www.bacpac.chori.org/porcine242.htm). For BAC library
  availability, please contact Pieter de Jong (pdejong@chori.org).
  Clones may be purchased from BACPAC Resources
  (http://BACPACorders.chori.org). This work was undertaken as part
  of the International Swine Genome Sequencing Consortium by
  University of Illinois at Urbana Champaign, USA with funds provided
  by Grant No. AG2002-34480-11828 from USDA-CSREES and
  AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
  Initiative)
  Plate: 431 row: D column: 15
  Seq primer: SP6
  Class: BAC ends.

FEATURES
  source
    1..538
      /organism="Sus scrofa"
      /mol_type="genomic DNA"
      /strain="four pigs (breed: 37.5% Yorks Landrace and 25%
      Meishan)"
      /db_xref="taxon:9823"
      /clone="RPCI44_431D15"
      /sex="male"
      /cell_type="blood"
      /clone_lib="RPCI-44"
      /notes="Vector: pPARAC2; Site 1: EcoRI; Site 2: EcoRI;
      porcine male BAC library produced by Pieter de Jong"

ORIGIN
  Query Match 66.7%; Score 20; DB 9; Length 538;
  Best Local Similarity 82.1%; Pred. No. 2.5e+02;
  Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

  1 AACAAATGATGTCGGGTGACATCTATG 28
    ||||| ||||| ||||| ||||| |||||
  2 AACAAATGATAACTGGGTGATATGATG 29

RESULT 15
AQ879634
LOCUS
DEFINITION
  HS_4816_A2_H10 T7A CIT Approved Human Genomic Sperm Library D Homo
  sapiens genomic clone Plate=4816 Col=20 Row=O, genomic survey
sequence.
ACCESSION
  AQ879634
VERSION
  AQ879634.1 GI:6311101
SOURCE
  GSS.
ORGANISM
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 542)
  Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
  Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
  Hood, L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL
  99380589
MEDLINE
  10449764
PUBMED
  Contact: Mahairas GG, Wallace JC, Hood L
  High Throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Clones may be purchased from Research Genetics (info@resgen.com).
  BAC end Web Server: http://www.htsc.washington.edu
  Plate: 4816 row: O column: 20
  Seq primer: T7
  Class: BAC ends
  High quality sequence stop: 542.

FEATURES
  source
    1..542
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="Plate=4816 Col=20 Row=O"
      /sex="male"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in
      E-Coli DH10B"

ORIGIN
  Query Match 66.7%; Score 20; DB 8; Length 542;
  Best Local Similarity 82.1%; Pred. No. 2.5e+02;
  Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

  2 ACAATGTATGTCGGGTGACATCTATGCA 29
    ||||| ||||| ||||| ||||| |||||
  157 ACAGTGTAGTAAGGTGTACACCTATGCA 130

RESULT 16
AQ525705/c
LOCUS
DEFINITION
  HS_5239_B1_C02 T7A RPCI-11 Human Male BAC Library Homo sapiens
  genomic clone Plate=815 Col=3 Row=F, genomic survey sequence.
ACCESSION
  AQ525705
VERSION
  AQ525705.1 GI:4773025
SOURCE
  GSS.
ORGANISM
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 605)
  Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
  Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
  Hood, L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL
  99380589
MEDLINE
  10449764
PUBMED
  Contact: Mahairas GG, Wallace JC, Hood L

```

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/bac.htm>) or from Resear h Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 815 row: F column: 3
Seq primer: T7
Class: BAC ends
High quality sequence stop: 605.
Location/Qualifiers
1. .605
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=815 Col=3 Row=F"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

ORIGIN
Query Match 66.7%; Score 20; DB 8; Length 605;
Best Local Similarity 82.1%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CAATGTATGTCGGGTGACATCTATGAC 30
|||||
Db 535 CAATGTATGTCGGGTGACATCTATGAC 508
|||||

RESULT 17
AQ856387
LOCUS
DEFINITION nbeb0003A23r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone nbeb0003A23r, genomic survey sequence.
ACCESSION AQ856387.1 GI:6206844
VERSION AQ856387
KEYWORDS GSS.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.
1. (bases 1 to 773)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 8
High quality sequence stop: 311.
Location/Qualifiers
1. .773
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"

FEATURES
source

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/bac.htm>) or from Resear h Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 815 row: F column: 3
Seq primer: T7
Class: BAC ends
High quality sequence stop: 605.
Location/Qualifiers
1. .605
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=815 Col=3 Row=F"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

ORIGIN
Query Match 66.7%; Score 20; DB 8; Length 605;
Best Local Similarity 82.1%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CAATGTATGTCGGGTGACATCTATGAC 30
|||||
Db 535 CAATGTATGTCGGGTGACATCTATGAC 508
|||||

RESULT 17
AQ856387
LOCUS
DEFINITION nbeb0003A23r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone nbeb0003A23r, genomic survey sequence.
ACCESSION AQ856387.1 GI:6206844
VERSION AQ856387
KEYWORDS GSS.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.
1. (bases 1 to 773)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 8
High quality sequence stop: 311.
Location/Qualifiers
1. .773
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"

FEATURES
source

/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbeb0003A23r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/notes="Vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN
Query Match 66.7%; Score 20; DB 8; Length 773;
Best Local Similarity 82.1%; Pred. No. 2.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGACATCTATG 28
|||||
Db 144 AACAAATGATGTCGGGTGACATCTACG 171
|||||

RESULT 18
CV480979/c
LOCUS
DEFINITION CV480979 858 bp mRNA linear EST 01-OCT-2004
AGENCOURT 33029658 NIH_ZGC_15 Danio rerio cDNA clone IMAGE:7440795
3', mRNA sequence.
CV480979
VERSION CV480979.1 GI:53737021
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1. (bases 1 to 858)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Will Talbot
cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LHAM15695 row: k column: 01
High quality sequence start: 81
High quality sequence stop: 743.
Location/Qualifiers
1. .858
/organism="Danio rerio"

FEATURES
source

/mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7440795"
 /tissue_type="whole embryos staged from 2-8 hr
 postfertilization, approximately 2500 embryos total"
 /lab_hosts="DH10B Tona"
 /clone_lib="NHG ZGC 15"
 /note="Organ: embryo; Vector: pME18S-FL3; Site 1: DraIII;
 Site 2: DraIII; 1st strand cDNA was primed with an
 oligo(dT) primer
 [GCGCTGAGACGGCTATGTCCTTTTTTTTTTTTTT];
 double-stranded cDNA was ligated to a DraIII adaptor
 [GGCUACUG], digested and directionally cloned into
 distinct DraIII sites of the pME18S-FL3. Library was size
 selected for 1.0 kb, with a average insert size of ~1.2kb.
 Library constructed by Yutaka Suzuki (University of Tokyo
 Institute of Medical Science). Custom primers recommended
 for sequencing: 5' end primer 5'-GGATGTGCTTTTACTTCA-3'
 and 3' end primer 5'-CGACCTGCAGCTCGACACA-3'. Note: This
 is a Zebrafish Gene Collection (ZGC) library"

ORIGIN

Query Match 66.7%; Score 20; DB 7; Length 859;
 Best Local Similarity 82.1%; Pred. No. 2.7e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 AACATGTATGTCGCGTGATCATCTATG 28

Db 766 AACATGTATGTCGCGTGATCATCTATG 739

RESULT 19

AZ204484/c

LOCUS

DEFINITION SP_0056_A1_A04_SP6E Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plates-56 Col=7 Row=A, genomic survey sequence.

ACCESSION AZ204484

VERSION AZ204484.1 GI:8399404

KEYWORDS

SOURCE Strongylocentrotus purpuratus

ORGANISM

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae; Strongylocentrotus. 1 (bases 1 to 916)
 Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R., Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and Hood,L.
 A sea urchin genome project: Sequence scan, virtual map, and additional resources

Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

20402566

10920195

Contact: Cameron, RA, Davidson, EH, Hood, L

Division of Biology 156-29

California Institute of Technology

Pasadena California 91125, USA

Tel: (626) 395-8421

Fax: (626) 793-3047

Email: acameron@caltech.edu

Plate: 56 row: A column: 7

Seq primer: SP6

Class: BAC ends

High quality sequence stop: 916.

Location/Qualifiers

1..916

/organism="Strongylocentrotus purpuratus"

/mol_type="genomic DNA"

/db_xref="taxon:7668"

/clone="Plate=56 Col=7 Row=A"

/clone_lib="Strongylocentrotus purpuratus, purple sea

urchin, sperm genomic BAC library"
 /note="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
 DH10B"

ORIGIN

Query Match 66.7%; Score 20; DB 8; Length 916;
 Best Local Similarity 82.1%; Pred. No. 2.7e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 2 ACAATGTATGTCGCGTGATCATCTATGA 29

Db 581 ACAATGTATGTCGCGTGATCATCTATGA 554

RESULT 20

LOCUS B12483/c

DEFINITION F26A16-T7 IGF Arabidopsis thaliana genomic clone F26A16, genomic survey sequence.

ACCESSION B12483

VERSION B12483.1 GI:2093769

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurooids II; Brassicales; Brassicaceae; Arabidopsais. 1 (bases 1 to 1330)

Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and Ecker,J.
 BAC End Sequences at ATGC
 Unpublished (1997)

JOURNAL

COMMENT Other GSSs: F26A16-T7.1, F26A16-T7.2, F26A16-SP6, F26A16-SP6.1

Contact: Ecker J.

Arabidopsis Thaliana Genome Center

University of Pennsylvania

Dept. of Biology, University of Pennsylvania, Philadelphia, PA

19104

Tel: 215-898-9384

Fax: 215-898-8780

Email: jecker@genome.bio.upenn.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 92

High quality sequence stop: 678.

FEATURES

source

1..1330

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/clone="F26A16"

/sex="hermaphrodite"

/clone_lib="IGF"

/note="Vector: BclBACII; Site 1: EcoRI; Site 2: EcoRI;

Produced by Thomas Altmann"

ORIGIN

Query Match 66.7%; Score 20; DB 8; Length 1330;
 Best Local Similarity 82.1%; Pred. No. 2.9e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 2 ACAATGTATGTCGCGTGATCATCTATGA 29

Db 136 ACAATGTATGTCGCGTGATCATCTATGA 109

RESULT 21

BZ091289

LOCUS

DEFINITION CH230-213G14-TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

CH230-213G14, genomic survey sequence.

ACCESSION BZ091289


```

VERSION      BZ091289.1  GI:23728348
KEYWORDS     GSS.
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM     Rattus norvegicus
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE    1 (bases 1 to 266)
AUTHORS     Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
              Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
              Riggs,F., de Jong,P. and Fraser,C.M.
TITLE       Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL     Unpublished (1999)
COMMENT     Other GSSs: CH230-213G14.TJ
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the rat BAC library CHORI-230
              (http://www.chori.org/bacpac/rat230.htm). For BAC library
              availability, please contact Pieter de Jong (pdejong@mail.cho.org).
              Clones may be purchased from BACPAC Resources
              (http://www.chori.org/bacpac/orering\_information.htm). BAC end
              page: http://www.tigr.org/tldb/bac\_ends/rat/bac\_end\_intro.html
              Plate: 213 row: G column: 14
              Seq primer: T7
              Class: BAC ends.

FEATURES     source
              Location/Qualifiers
                1..266
                  /organism="Rattus norvegicus"
                  /mol_type="genomic DNA"
                  /strain="BN/SENHed/MCW"
                  /db_xref="taxon:10116"
                  /clone="CH230-213G14"
                  /sex="Female"
                  /cell_type="Brain"
                  /clone_lib="CHORI-230 Segment 1"
                  /notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                  CHORI-230 Rat (BN/SENHed/MCW) BAC library produced by
                  Pieter de Jong"

ORIGIN
Query Match      65.3%; Score 19.6; DB 8; Length 266;
Best Local Similarity 84.6%; Pred. No. 3.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTACATCTATGAC 30
    ||||| ||||| ||||| ||||| |||||
Db 148 ATGTGTGTCGGGTACATATATGTC 173

RESULT 22
LOCUS      AA751822
DEFINITION 96AS0663 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
            (indica cultivar-group) cDNA clone 96AS0663, mRNA sequence.
ACCESSION  AA751822
VERSION     AA751822.1  GI:2798528
KEYWORDS   EST.
SOURCE     Oryza sativa (indica cultivar-group)
ORGANISM   Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 380)
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C. and Eun,M.Y.
TITLE     Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL   Unpublished (1998)

COMMENT    Contact: Eun M.Y.
            Department of Cytogenetics
            National Inst. of Agri. Sci. and Tech, RDA
            Suwon, Kyunggido, Korea
            Tel: 82 331 290 0301
            Fax: 82 331 290 0307
            Email: myeun@sun20.asti.re.kr
            Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
            University, Yongin, Korea. 449-728 bhnahm@bioserver.myongji.ac.kr
            Seq primer: M13 Reverse Primer.

FEATURES     source
              Location/Qualifiers
                1..380
                  /organism="Oryza sativa (indica cultivar-group)"
                  /mol_type="mRNA"
                  /cultivar="Milyang 23"
                  /db_xref="taxon:39946"
                  /clone="96AS0663"
                  /tissue_type="Immature Seed"
                  /dev_stage="5 days after pollination"
                  /lab_host="E. coli SOLR"
                  /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
                  /notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
                  XhoI; Directional cDNA library inserted into lambda ZAPII
                  vector at 5' end with EcoRI and 3' end with Xho I site."

ORIGIN
Query Match      65.3%; Score 19.6; DB 1; Length 380;
Best Local Similarity 78.6%; Pred. No. 3.6e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACATGTATGTCGGGTACATCTATG 28
    ||||| ||||| ||||| ||||| |||||
Db 327 AAAAATGCATCTTNANGTACATCTATG 354

RESULT 23
LOCUS      BH121074
DEFINITION RPCI-24-230E9.TV RPCI-24 Mus musculus genomic clone RPCI-24-230E9,
            genomic survey sequence.
ACCESSION  BH121074
VERSION     BH121074.1  GI:14964586
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 385)
AUTHORS   Zhao,S., Niernan,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Teegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
TITLE     Mouse BAC End Sequences from Library RPCI-24
JOURNAL   Unpublished (1999)
COMMENT    Other GSSs: RPCI-24-230E9.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://www.tigr.org/tldb/bac\_ends/mouse/bac\_end\_intro.html
            Plate: 230 row: E column: 9
            Seq primer: T7
            Class: BAC ends.

FEATURES     source
              Location/Qualifiers
                1..385
                  /organism="Mus musculus"
                  /mol_type="genomic DNA"

```

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-230E9"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1, Site 1: BamH1, Site 2: BamH1,
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN
Query Match 65.3%; Score 19.6; DB 8; Length 385;
Best Local Similarity 84.6%; Pred. No. 3.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGTACATCTAT 27
|||||
Db 166 ACAGTGTATGTCCTGTCTACATTTAT 191

RESULT 24
CF889471
LOCUS TcTR-272 TcTR Trypanosoma cruzi cDNA clone 02e8 5', mRNA linear EST 04-MAR-2004
DEFINITION TcTR-272 TcTR Trypanosoma cruzi cDNA clone 02e8 5', mRNA sequence.
ACCESSION CF889471
VERSION 1
KEYWORDS EST.
SOURCE Trypanosoma cruzi
ORGANISM Trypanosoma cruzi
Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
1 (bases 1 to 396)
Aguero, P., Ben Abdellah, K., Tekiel, V., Sanchez, D.O. and Gonzalez, A.
Generation and analysis of expressed sequence tags from Trypanosoma
cruzi trypanastigote and amastigote cDNA libraries
Mol. Biochem. Parasitol. 136 (2), 221-225 (2004)
Contact: Sanchez DO
Genomics and Bioinformatics
Instituto de Investigaciones Biotecnologicas
Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos
Aires, Argentina
Tel: (54-11) 4580/7255/7
Fax: (54-11) 4752-9639
Email: dsanchez@iib.unsam.edu.ar
Sequences were basecalled with phred and vector was masked with
crossmatch (see http://www.phrap.org). Sequences were then trimmed
from both ends to remove low quality bases and masked vector.
Plate: 02 row: e column: 8
Seq primer: T7.

FEATURES
Location/Qualifiers
1..396
/organism="Trypanosoma cruzi"
/mol_type="mRNA"
/db_xref="taxon:5693"
/clone="02e8"
/dev_stage="trypanastigote"
/clone_lib="TcTR"

Note="Vector: pBluescript KS+; Site 1: EcoRI; Site 2:
NotI; PolyA(+) RNA was prepared using the QuickPrep Micro
RNA Purification Kit (Amersham Pharmacia Biotech).
Synthesis of the cDNA first strand was done with
Superscript II reverse transcriptase (Gibco BRL) and an
oligo-dT-not primer (5'-CTCGGCCGCT(18)-3'). Synthesis of
the cDNA second strand was performed using the Klenow
fragment of DNA polymerase with a spliced leader (SL)
primer (5'-GATCAGTTCGTGA-3'). After methylation with
EcoRI methylase, phosphorylated EcoRI linkers
(5'-ACCGAATTCTGT-3') were ligated to the cDNA. The
resulting cDNA mixture was then digested with NotI and
EcoRI restriction enzymes, subjected to size fractionation
on SizeSep400 Spun Columns (Pharmacia) to select fragments

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 09:43:25 ; Search time 90 Seconds
(without alignments)
545.426 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgtatgcgtgtacatcatgac 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.6	78.7	2042	2	US-08-911-434A-2
2	19	63.3	198632	4	US-09-949-016-12781
3	19	63.3	198637	4	US-09-949-016-17393
C 4	18.6	62.0	601	4	US-09-949-016-103581
5	18.6	62.0	2152	4	US-09-949-632-3
C 6	18.6	62.0	256287	4	US-09-949-016-14608
C 7	18.4	61.3	40512	4	US-09-949-016-16612
C 8	18	60.0	6439	4	US-09-902-540-813
9	17.8	59.3	438	3	US-08-651-155B-134
10	17.8	59.3	40742	4	US-09-194-036B-134
11	17.8	59.3	908	3	US-09-457-046B-15
12	17.8	59.3	908	4	US-09-866-570B-15
13	17.8	59.3	1686	4	US-09-902-540-9105
14	17.8	59.3	10318	4	US-09-902-540-973
15	17.8	59.3	19698	4	US-09-949-016-14079
C 16	17.8	59.3	40742	4	US-09-949-016-11751
C 17	17.8	59.3	40747	4	US-09-949-016-13097
C 18	17.8	59.3	70088	4	US-09-949-016-16845
19	17.8	59.3	123463	4	US-09-949-016-17078
20	17.8	59.3	235452	4	US-09-949-016-13675
21	17.6	58.7	601	4	US-09-949-016-32764
22	17.6	58.7	601	4	US-09-949-016-32823
23	17.6	58.7	601	4	US-09-949-016-50681
24	17.6	58.7	601	4	US-09-949-016-50740
C 25	17.6	58.7	27702	4	US-09-949-016-11795
C 26	17.6	58.7	77851	4	US-09-949-016-12508
C 27	17.6	58.7	77867	4	US-09-949-016-13211

C 28	17.6	58.7	77867	4	US-09-949-016-13212	Sequence 13212, A
C 29	17.6	58.7	77940	4	US-09-949-016-12509	Sequence 12509, A
C 30	17.6	58.7	312957	4	US-09-949-001-31	Sequence 31, Appl
C 31	17.6	58.7	312972	4	US-09-949-001-34	Sequence 34, Appl
C 32	17.4	58.0	198	4	US-09-248-796A-7260	Sequence 7260, Ap
C 33	17.4	58.0	1507	3	US-08-605-150A-15	Sequence 15, Appl
C 34	17.4	58.0	2182	3	US-09-221-017B-916	Sequence 916, Appl
C 35	17.4	58.0	2662	3	US-08-750-357-8	Sequence 8, Appl
C 36	17.4	58.0	3848	3	US-09-112-096-28	Sequence 28, Appl
C 37	17.4	58.0	5668	3	US-09-112-096-14	Sequence 14, Appl
C 38	17.4	58.0	5668	4	US-09-636-215-777	Sequence 777, App
C 39	17.4	58.0	5668	4	US-09-685-166A-777	Sequence 777, App
C 40	17.4	58.0	5668	4	US-09-679-426-777	Sequence 777, App
C 41	17.4	58.0	5668	4	US-09-759-143-777	Sequence 777, App
C 42	17.4	58.0	5668	4	US-09-651-236-777	Sequence 777, App
C 43	17.4	58.0	48974	3	US-08-920-422-17	Sequence 17, Appl
C 44	17.2	57.3	351	4	US-09-328-352-182	Sequence 182, App
C 45	17.2	57.3	601	4	US-09-949-016-71655	Sequence 71655, A
C 46	17.2	57.3	601	4	US-09-949-016-71656	Sequence 71656, A
C 47	17.2	57.3	601	4	US-09-949-016-71728	Sequence 71728, A
C 48	17.2	57.3	601	4	US-09-949-016-71729	Sequence 71729, A
C 49	17.2	57.3	601	4	US-09-949-016-71801	Sequence 71801, A
C 50	17.2	57.3	601	4	US-09-949-016-71802	Sequence 71802, A
C 51	17.2	57.3	601	4	US-09-949-016-71874	Sequence 71874, A
C 52	17.2	57.3	601	4	US-09-949-016-71875	Sequence 71875, A
C 53	17.2	57.3	601	4	US-09-949-016-71947	Sequence 71947, A
C 54	17.2	57.3	601	4	US-09-949-016-71948	Sequence 71948, A
C 55	17.2	57.3	601	4	US-09-949-016-72020	Sequence 72020, A
C 56	17.2	57.3	601	4	US-09-949-016-72021	Sequence 72021, A
C 57	17.2	57.3	601	4	US-09-949-016-72093	Sequence 72093, A
C 58	17.2	57.3	601	4	US-09-949-016-72094	Sequence 72094, A
C 59	17.2	57.3	601	4	US-09-949-016-72166	Sequence 72166, A
C 60	17.2	57.3	601	4	US-09-949-016-72167	Sequence 72167, A
C 61	17.2	57.3	601	4	US-09-949-016-72239	Sequence 72239, A
C 62	17.2	57.3	601	4	US-09-949-016-72240	Sequence 72240, A
C 63	17.2	57.3	601	4	US-09-949-016-72312	Sequence 72312, A
C 64	17.2	57.3	601	4	US-09-949-016-72313	Sequence 72313, A
65	17.2	57.3	5474	1	US-08-137-252-2	Sequence 2, Appl
66	17.2	57.3	11663	1	US-08-446-932-1	Sequence 1, Appl
67	17.2	57.3	11663	1	US-08-801-263A-1	Sequence 1, Appl
68	17.2	57.3	11663	1	US-08-801-263A-7	Sequence 7, Appl
69	17.2	57.3	11663	3	US-09-102-248-1	Sequence 1, Appl
70	17.2	57.3	11663	3	US-09-102-248-7	Sequence 7, Appl
71	17.2	57.3	11663	3	US-09-367-764-1	Sequence 1, Appl
72	17.2	57.3	11663	4	US-09-367-764-7	Sequence 7, Appl
C 73	17.2	57.3	86945	4	US-09-949-016-13849	Sequence 13849, A
C 74	17.2	57.3	86945	4	US-09-949-016-13850	Sequence 13850, A
C 75	17.2	57.3	86945	4	US-09-949-016-13851	Sequence 13851, A
C 76	17.2	57.3	86945	4	US-09-949-016-13852	Sequence 13852, A
C 77	17.2	57.3	86945	4	US-09-949-016-13853	Sequence 13853, A
C 78	17.2	57.3	86945	4	US-09-949-016-13854	Sequence 13854, A
C 79	17.2	57.3	86945	4	US-09-949-016-13855	Sequence 13855, A
C 80	17.2	57.3	86945	4	US-09-949-016-13856	Sequence 13856, A
C 81	17.2	57.3	86945	4	US-09-949-016-13857	Sequence 13857, A
C 82	17.2	57.3	86945	4	US-09-949-016-13858	Sequence 13858, A
C 83	17.2	57.3	154023	4	US-09-949-016-17057	Sequence 17057, A
C 84	17.2	57.3	312957	4	US-09-949-001-31	Sequence 31, Appl
85	17.2	57.3	312972	4	US-09-949-001-34	Sequence 34, Appl
86	17	56.7	601	4	US-09-949-016-205744	Sequence 205744, A
87	17	56.7	601	4	US-09-949-016-205745	Sequence 205745, A
88	17	56.7	601	4	US-09-949-016-205746	Sequence 205746, A
89	17	56.7	745	3	US-09-270-767-12913	Sequence 12913, A
90	17	56.7	11580	3	US-09-334-220-4	Sequence 4, Appl
C 91	17	56.7	48940	4	US-09-949-016-16402	Sequence 16402, A
C 92	17	56.7	152524	4	US-09-949-016-12683	Sequence 12683, A
C 93	17	56.7	152524	4	US-09-949-016-13194	Sequence 13194, A
94	17	56.7	254405	4	US-09-949-016-14381	Sequence 14381, A
95	17	56.7	271134	4	US-09-949-016-12705	Sequence 12705, A
96	17	56.7	305431	1	US-09-949-016-17550	Sequence 17550, A
97	16.8	56.0	275	1	US-07-789-919A-1	Sequence 1, Appl
98	16.8	56.0	275	1	US-08-209-846A-1	Sequence 1, Appl
99	16.8	56.0	275	2	US-08-472-809B-1	Sequence 1, Appl
100	16.8	56.0	275	3	US-08-438-265-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911.434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Goshun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..2042
US-08-911-434A-2

Query Match 78.7%; Score 23.6; DB 2; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.27;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
|||||
Db 1737 AACAACTTTGTCGGGTATATTATGAC 1766

RESULT 2

US-09-949-016-12781
; Sequence 12781, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12781
; LENGTH: 198632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12781

Query Match 63.3%; Score 19; DB 4; Length 198632;
Best Local Similarity 81.5%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 AATGTATGTCGGGTGTACATCTATGAC 30
|||||
Db 142458 AATGTGTGTAGATGATCATCTGTGAC 142484

RESULT 3

US-09-949-016-17393
; Sequence 17393, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17393
; LENGTH: 198637
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17393

Query Match 63.3%; Score 19; DB 4; Length 198637;
Best Local Similarity 81.5%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 AATGTATGTCGGGTGTACATCTATGAC 30
|||||
Db 142458 AATGTGTGTAGATGATCATCTGTGAC 142484

RESULT 4

US-09-949-016-103581/c
; Sequence 103581, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

RESOLV 8
US-09-902-540-813/c
; Sequence 813, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 813
; LENGTH: 6439
; TYPE: DNA
; ORGANISM: Myxococcus xanthus

```

US-09-902-540-813

Query Match 60.0%; Score 18; DB 4; Length 6439;
Best Local Similarity 80.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGTACATCATGAC 30
||||| ||||| ||||| ||||| |||||
Db 5287 ATGTATGGCCGCTGTGCATCAATGGC 5262

RESULT 9

US-08-651-155B-134
; Sequence 134, Application US/08651155B
; Patent No. 6365401
; GENERAL INFORMATION:
; APPLICANT: Mahan Dr., Michael J.
; APPLICANT: Conner Mr., Christopher P.
; APPLICANT: Hiethoff Mr., Douglas M.
; TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION
; OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
; INFECTION
; NUMBER OF SEQUENCES: 255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chrisman, Bynum & Johnson, P.C.
; STREET: 1900 Fifteenth Street
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80302
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,155B
; FILING DATE: 17-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Petersen Mr., Steven C.
; REGISTRATION NUMBER: 36,238
; REFERENCE/DOCKET NUMBER: 17060.1
; TELEPHONE: 303/546-1300
; TELEFAX: 303/449-5426
; TELEX: ABA1475
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-651-155B-134

Query Match 59.3%; Score 17.8; DB 3; Length 438;
Best Local Similarity 75.9%; Pred. No. 1e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGTACATCATGAC 30
||||| ||||| ||||| ||||| |||||
Db 149 ATAATGAATGGCGGTGTACACCCATGAC 177

RESULT 10

US-09-194-036B-134
; Sequence 134, Application US/09194036B
; Patent No. 6548246
; GENERAL INFORMATION:
; APPLICANT: Mahan, Michael J.

Conner, Christopher P.
Hiethoff, Douglas M.
TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION
OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST
INFECTION
NUMBER OF SEQUENCES: 255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Mountain View
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/194,036B
FILING DATE: 17-No. 6548246-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/08208
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 08/651,155
FILING DATE: 1996-05-17
ATTORNEY/AGENT INFORMATION:
NAME: Shantanu Basu
REGISTRATION NUMBER: 43,318
REFERENCE/DOCKET NUMBER: 220002060601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5995
TELEFAX: (650) 494-0792
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 438 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: YES
ORIGINAL SOURCE:
ORGANISM: DNA (other)
SEQUENCE DESCRIPTION: SEQ ID NO: 134:
US-09-194-036B-134
Query Match 59.3%; Score 17.8; DB 4; Length 438;
Best Local Similarity 75.9%; Pred. No. 1e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 2 ACAATGATGTCGGGTGTACATCATGAC 30
||||| ||||| ||||| ||||| |||||
Db 149 ATAATGAATGGCGGTGTACACCCATGAC 177
RESULT 11
US-09-457-046B-15
; Sequence 15, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Taxus cuspidata
US-09-457-046B-15

Sequence 375, Application 00/0550231
: Patent No. 6833447

2 ACAATGTATGTCGGGTACATCTATGAC 30


```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11751
; LENGTH: 40742
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11751

Query Match          59.3%; Score 17.8; DB 4; Length 40742;
Best Local Similarity 75.9%; Pred. No. 2.1e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AACAACTGATGTCGGGTGTACATCTATGA 29
Db 10725 AACAAACCATGTCGTGTACTTCTCTTA 10697

RESULT 17
US-09-949-016-13097/c
; Sequence 13097, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13097
; LENGTH: 40747
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13097

Query Match          59.3%; Score 17.8; DB 4; Length 40747;
Best Local Similarity 75.9%; Pred. No. 2.1e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AACAACTGATGTCGGGTGTACATCTATGA 29
Db 10725 AACAAACCATGTCGTGTACTTCTCTTA 10697

RESULT 18
US-09-949-016-16845/c
; Sequence 16845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16845
; LENGTH: 70088
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16845

Query Match          59.3%; Score 17.8; DB 4; Length 70088;
Best Local Similarity 75.9%; Pred. No. 2.3e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 AACATGTATGTCGGGTGTACATCTATGAC 30
Db 70027 ACAATGTGTCTCTGTACTCTCAAGCAC 69999

RESULT 19
US-09-949-016-17078
; Sequence 17078, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17078
; LENGTH: 123463
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(123463)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17078

Query Match          59.3%; Score 17.8; DB 4; Length 123463;
Best Local Similarity 75.9%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AACAACTGATGTCGGGTGTACATCTATGA 29
Db 22396 AACAAATAATACCTGTGTACATCTGTGA 22424

RESULT 20
US-09-949-016-13675
; Sequence 13675, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

RESULT 22
US-09-949-016-32823
; Sequence 32823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN K
; TITLE OF INVENTION: WITH HUMAN DISEAS
; FILE REFERENCE: C001307
; CURRENT APPLICATION NUMBER: US/09/949,
; CURRENT FILING DATE: 2000-04-14

RESULT 24
US-09-949-016-50740
; Sequence 50740, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN K
; TITLE OF INVENTION: WITH HUMAN DISEAS
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50740
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-50740

Query Match      58.7%; Score 17.6; DB 4; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AACAACTATGTCGGGTGTACATC 24
Db      193 AATAATATATCTCCAGTGTACATC 216

RESULT 25
US-09-949-016-11795/c
; Sequence 11795, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11795
; LENGTH: 27702
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11795

Query Match      58.7%; Score 17.6; DB 4; Length 27702;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      3 CAATGTATGTCGGGTGTACATCTA 26
Db      18205 CAATGTTTCCTCCGGTGTCTCTA 18182

Search completed: October 28, 2005, 11:18:22
Job time : 94 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 10:52:30 ; Search time 425 Seconds
(without alignments)
582.988 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgtatgtccggtacatcatgac 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19.6	65.3	380	20	US-10-437-963-26247, A
C 2	19.6	65.3	1114	13	US-09-925-065A-708892
C 3	19.4	64.7	551	13	US-09-925-065A-193679
C 4	19.4	64.7	374849	14	US-10-087-192-1627
C 5	19.2	64.0	42772	14	US-10-087-192-1903

C	6	19.2	64.0	2731748	20	US-10-297-465A-1	Sequence 1, Appli
C	7	19	63.3	512	20	US-10-021-323-6008	Sequence 6008, Ap
C	8	19	63.3	545	13	US-09-925-065A-163111	Sequence 163111,
C	9	19	63.3	551	13	US-09-925-065A-193680	Sequence 193680,
10	19	63.3	35042	24	US-10-915-740A-2	Sequence 2, Appli	
C	11	19	63.3	50000	15	US-10-152-724A-22	Sequence 22, Appli
C	12	19	63.3	1601042	14	US-10-027-632-59064	Sequence 59064, A
C	13	19	63.3	1601042	18	US-10-027-632-59064	Sequence 59064, A
C	14	19	63.3	2242716	24	US-10-915-740A-1068	Sequence 1068, Ap
C	15	18.8	62.7	1630	9	US-09-822-830A-444	Sequence 444, App
C	16	18.6	62.0	665	14	US-10-027-632-8033	Sequence 8033, Ap
C	17	18.6	62.0	665	18	US-10-027-632-8033	Sequence 8033, Ap
C	18	18.6	62.0	2152	17	US-10-372-095-3	Sequence 3, Appli
C	19	18.6	62.0	90442	15	US-10-105-637-1	Sequence 1, Appli
C	20	18.6	62.0	90442	18	US-10-034-650-43	Sequence 43, Appli
C	21	18.4	61.3	255	19	US-10-424-599-132779	Sequence 132779,
C	22	18.4	61.3	650	14	US-10-027-632-320106	Sequence 320106,
C	23	18.4	61.3	650	18	US-10-027-632-320106	Sequence 320106,
C	24	18.4	61.3	658	13	US-09-925-065A-765698	Sequence 765698,
C	25	18.4	61.3	662	14	US-10-027-632-84202	Sequence 84202, A
C	26	18.4	61.3	662	18	US-10-027-632-84202	Sequence 84202, A
C	27	18.4	61.3	3177	26	US-11-097-143-40196	Sequence 40196, A
C	28	18.4	61.3	4354	26	US-11-097-143-13798	Sequence 13798, A
C	29	18.4	61.3	5460	26	US-11-097-143-40195	Sequence 40195, A
C	30	18.4	61.3	188017	14	US-10-087-192-1951	Sequence 1951, Ap
C	31	18.2	60.7	4919	26	US-11-097-143-20047	Sequence 20047, A
C	32	18.2	60.7	40050	19	US-10-052-482-91	Sequence 91, Appli
C	33	18.2	60.7	310122	21	US-10-417-375-1	Sequence 1, Appli
C	34	18	60.0	287	19	US-10-621-901-1833	Sequence 1833, Ap
C	35	18	60.0	287	19	US-10-621-901-1859	Sequence 1859, Ap
C	36	18	60.0	386	19	US-10-621-901-1927	Sequence 1927, Ap
C	37	18	60.0	401	9	US-09-795-668-1084	Sequence 1084, Ap
C	38	18	60.0	401	9	US-09-795-668-1084	Sequence 1084, Ap
C	39	18	60.0	401	9	US-09-946-807-1084	Sequence 1084, Ap
C	40	18	60.0	401	24	US-10-995-011-1084	Sequence 1084, Ap
C	41	18	60.0	448	19	US-10-621-901-1276	Sequence 1276, Ap
C	42	18	60.0	546	13	US-09-925-065A-620304	Sequence 620304, A
C	43	18	60.0	599	24	US-10-972-079-57689	Sequence 57689, A
C	44	18	60.0	603	13	US-09-925-065A-744005	Sequence 744005,
C	45	18	60.0	620	21	US-10-425-115-32472	Sequence 32472, A
C	46	18	60.0	649	19	US-10-424-599-103647	Sequence 103647,
C	47	18	60.0	2194	20	US-10-437-963-21442	Sequence 21442, A
C	48	18	60.0	2880	19	US-10-424-599-74550	Sequence 74550, A
C	49	18	60.0	42007	18	US-10-085-117-337	Sequence 337, App
C	50	18	60.0	83709	21	US-10-723-860-2784	Sequence 2784, Ap
C	51	18	60.0	1503841	9	US-09-795-668-1	Sequence 1, Appli
C	52	18	60.0	1503841	9	US-09-795-668-1	Sequence 1, Appli
C	53	18	60.0	1503841	9	US-09-946-807-1	Sequence 1, Appli
C	54	18	60.0	1503841	24	US-10-995-011-1	Sequence 1, Appli
C	55	17.8	59.3	141	14	US-10-033-528-1890	Sequence 1890, Ap
C	56	17.8	59.3	141	17	US-10-099-926-1890	Sequence 1890, Ap
C	57	17.8	59.3	421	20	US-10-437-963-89493	Sequence 89493, A
C	58	17.8	59.3	428	15	US-10-198-846-922	Sequence 922, App
C	59	17.8	59.3	440	24	US-10-450-763-2874	Sequence 2874, Ap
C	60	17.8	59.3	500	13	US-09-918-995-23954	Sequence 23954, A
C	61	17.8	59.3	540	13	US-09-925-065A-222468	Sequence 222468,
C	62	17.8	59.3	540	13	US-09-925-065A-222469	Sequence 222469,
C	63	17.8	59.3	540	13	US-09-925-065A-222470	Sequence 222470,
C	64	17.8	59.3	545	13	US-09-925-065A-163112	Sequence 163112,
C	65	17.8	59.3	578	13	US-09-925-065A-129747	Sequence 129747,
C	66	17.8	59.3	662	15	US-10-198-846-829	Sequence 829, App
C	67	17.8	59.3	719	13	US-09-925-065A-8079	Sequence 8079, Ap
C	68	17.8	59.3	719	13	US-09-925-065A-8080	Sequence 8080, Ap
C	69	17.8	59.3	791	13	US-09-925-065A-31609	Sequence 31609, A
C	70	17.8	59.3	791	13	US-09-925-065A-31610	Sequence 31610, A
C	71	17.8	59.3	881	15	US-10-198-846-10611	Sequence 10611, A
C	72	17.8	59.3	908	9	US-09-866-572A-15	Sequence 15, Appli
C	73	17.8	59.3	908	16	US-10-166-984-15	Sequence 15, Appli
C	74	17.8	59.3	908	18	US-10-166-984-15	Sequence 15, Appli
C	75	17.8	59.3	1513	21	US-10-739-930-1153	Sequence 1153, Ap
C	76	17.8	59.3	1823	14	US-10-027-632-99290	Sequence 99290, A
C	77	17.8	59.3	1823	14	US-10-027-632-99290	Sequence 99290, A
C	78	17.8	59.3	1823	18	US-10-027-632-99290	Sequence 99290, A

Publication No. US20020182586A1
 GENERAL INFORMATION:
 APPLICANT: Morris, David W.
 APPLICANT: Engelhard, Eric K.
 TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 TITLE OF INVENTION: CANCER
 FILE REFERENCE: 529452000122
 CURRENT APPLICATION NUMBER: US/10/087,192
 CURRENT FILING DATE: 2002-03-01
 PRIOR APPLICATION NUMBER: US 09/747,377
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: US 09/798,586
 PRIOR FILING DATE: 2001-03-02
 NUMBER OF SEQ ID NOS: 2059
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1627
 LENGTH: 374849
 TYPE: DNA
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)...(374849)
 OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-1627

Query Match 64.7%; Score 19.4; DB 14; Length 374849;
 Best Local Similarity 79.3%; Pred. No. 2.5e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGA 29
 Db 6156 AACAAATGATGTCAGTGACAGGTATTA 6184

RESULT 5

US-10-087-192-1903/c
 Sequence 1903, Application US/10087192
 Publication No. US20020182586A1
 GENERAL INFORMATION:
 APPLICANT: Morris, David W.
 APPLICANT: Engelhard, Eric K.
 TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 TITLE OF INVENTION: CANCER
 FILE REFERENCE: 529452000122
 CURRENT APPLICATION NUMBER: US/10/087,192
 CURRENT FILING DATE: 2002-03-01
 PRIOR APPLICATION NUMBER: US 09/747,377
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: US 09/798,586
 PRIOR FILING DATE: 2001-03-02
 NUMBER OF SEQ ID NOS: 2059
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1903
 LENGTH: 42772
 TYPE: DNA
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)...(42772)
 OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-1903

Query Match 64.0%; Score 19.2; DB 14; Length 42772;
 Best Local Similarity 87.5%; Pred. No. 2.2e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGTACATCTATG 28
 Db 15706 ATGTATGTCGGGTGTACATCTATG 15683

RESULT 6

US-10-297-465A-1/c

Sequence 1, Application US/10297465A
 Publication No. US20040142413A1
 GENERAL INFORMATION:
 APPLICANT: Simpson, Andrew
 APPLICANT: Reinach, Fernando
 APPLICANT: Setubal, Joao
 APPLICANT: Medanis, Joao
 APPLICANT: Arruda, Paulo
 TITLE OF INVENTION: Isolated Genome of Xylella Fastidiosa and Uses Thereof
 FILE REFERENCE: FAPESP 202 US (10213376)
 CURRENT APPLICATION NUMBER: US/10/297,465A
 CURRENT FILING DATE: 2001-06-07
 PRIOR APPLICATION NUMBER: PCT/IB01/01618
 PRIOR FILING DATE: 2001-06-07
 PRIOR APPLICATION NUMBER: 60/209,906
 PRIOR FILING DATE: 2001-06-17
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 1
 LENGTH: 2731748
 TYPE: DNA
 ORGANISM: Xylella fastidiosa
 US-10-297-465A-1

Query Match 64.0%; Score 19.2; DB 20; Length 2731748;
 Best Local Similarity 87.5%; Pred. No. 3.9e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATC 24
 Db 1072833 AACAAATGATGTCGGGTGTACATC 1072810

RESULT 7

US-10-021-323-6008
 Sequence 6008, Application US/10021323
 Publication No. US20040123340A1
 GENERAL INFORMATION:
 APPLICANT: Deikman, Jill
 APPLICANT: Feng, Paul C.C.
 APPLICANT: Fincher, Karen L.
 APPLICANT: Ziegler, Todd E.
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants
 FILE REFERENCE: 38-21(52274)B
 CURRENT APPLICATION NUMBER: US/10/021,323
 CURRENT FILING DATE: 2001-12-12
 PRIOR APPLICATION NUMBER: US 60/255, 619
 PRIOR FILING DATE: 2000-12-14
 NUMBER OF SEQ ID NOS: 17880
 SEQ ID NO 6008
 LENGTH: 512
 TYPE: DNA
 ORGANISM: Gossypium hirsutum
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB3828-012-Q1-N6-C3
 US-10-021-323-6008

Query Match 63.3%; Score 19; DB 20; Length 512;
 Best Local Similarity 81.5%; Pred. No. 1.4e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CAATGATGTCGGGTGTACATCTATGA 29
 Db 394 CAATCCATTTCTGGGTGTACATCTCTGA 420

RESULT 8

US-09-925-065A-163111/c
 Sequence 163111, Application US/09925065A
 Publication No. US20050228172A9
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163111
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-163111

Query Match 63.3%; Score 19; DB 13; Length 545;
Best Local Similarity 75.9%; Pred. No. 1.5e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGATCATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 141 AACCCWATATGTCGGGTGCTCTTCTATGA 113

RESULT 9

US-09-925-065A-193680
; Sequence 193680, Application US/09925065A
; Publication No. US20050228172A9

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193680
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-193680

Query Match 63.3%; Score 19; DB 13; Length 551;
Best Local Similarity 75.9%; Pred. No. 1.5e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGATCATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 270 AACCCWATATGTCGGGTGCTCTTCTATGA 298

RESULT 10

US-10-915-740A-2
; Sequence 2, Application US/10915740A

; Publication No. US20050191316A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tettelin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Masignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroa
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scarlato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; PRIOR FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 35042
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; US-10-915-740A-2

Query Match 63.3%; Score 19; DB 24; Length 35042;
Best Local Similarity 81.5%; Pred. No. 2.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGATCATCTAT 27
||||| ||||| ||||| ||||| |||||
Db 8712 AACAAATGATGTCATCTAT 8738

RESULT 11

US-10-152-724A-22
; Sequence 22, Application US/10152724A
; Publication No. US20030082714A1
; GENERAL INFORMATION:
; APPLICANT: LITTLE, Melissa
; APPLICANT: HOLMES, Gregory
; APPLICANT: KOLLE, Gabriel
; APPLICANT: YAMADA, Toshiya
; APPLICANT: GEORGAS, Kylie
; APPLICANT: WILKINSON, Lorine
; TITLE OF INVENTION: No. US20030082714A1el Nucleic Acid and Polypeptide
; FILE REFERENCE: P22378
; CURRENT APPLICATION NUMBER: US/10/152,724A
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: Australian App No. US20030082714A1PQ 4348
; PRIOR FILING DATE: 1999-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-152-724A-22

Query Match 63.3%; Score 19; DB 15; Length 50000;
Best Local Similarity 81.5%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 AATGTATGTCGGGTACATCTATGAC 30
 ||||| ||| | ||||| |||||
 Db 33196 AATGTGTAGATGTACATCTGTGAC 33222

RESULT 12

US-10-027-632-59064/c
 ; Sequence 59064, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 59064
 ; LENGTH: 1601042
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1601042)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-027-632-59064

Query Match 63.3%; Score 19; DB 14; Length 1601042;
 Best Local Similarity 81.5%; Pred. No. 4.5e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ACAATGTATGTCGGGTACATCTATG 28
 ||||| ||| | ||||| |||||
 Db 707259 ACAATGTGTGCGGTCAAAATCTATG 707233

RESULT 13

US-10-027-632-59064/c
 ; Sequence 59064, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 59064
 ; LENGTH: 1601042
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1601042)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-027-632-59064

Query Match 63.3%; Score 19; DB 18; Length 1601042;
 Best Local Similarity 81.5%; Pred. No. 4.5e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ACAATGTATGTCGGGTACATCTATG 28
 ||||| ||| | ||||| |||||
 Db 707259 ACAATGTGTGCGGTCAAAATCTATG 707233

RESULT 14

US-10-915-740A-1068
 ; Sequence 1068, Application US/10915740A
 ; Publication No. US20050191316A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frazer, Claire M.
 ; APPLICANT: Hickey, Erin
 ; APPLICANT: Peterson, Jeremy
 ; APPLICANT: Tettelin, Herve
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: Massignani, Vega
 ; APPLICANT: Galeotti, Cesira
 ; APPLICANT: Mora, Manroa
 ; APPLICANT: Ratti, Giulio
 ; APPLICANT: Scarselli, Maria
 ; APPLICANT: Scariato, Vincenzo
 ; APPLICANT: Rappuoli, Rino
 ; APPLICANT: Pizza, Mariagrazia
 ; APPLICANT: Grandi, Guido
 ; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
 ; FILE REFERENCE: 002441.00090
 ; CURRENT APPLICATION NUMBER: US/10/915,740A
 ; CURRENT FILING DATE: 2004-08-11
 ; PRIOR APPLICATION NUMBER: 09/806,866
 ; PRIOR FILING DATE: 1999-10-08
 ; PRIOR APPLICATION NUMBER: USSN 60/103,794
 ; PRIOR FILING DATE: 1998-10-09
 ; PRIOR APPLICATION NUMBER: USSN 60/132,068
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/25373
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 1068
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1068
 ; LENGTH: 2242716
 ; TYPE: DNA
 ; ORGANISM: Neisseria meningitidis
 US-10-915-740A-1068

Query Match 63.3%; Score 19; DB 24; Length 2242716;
 Best Local Similarity 81.5%; Pred. No. 4.7e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAAATGTATGTCGGGTACATCTAT 27
 ||||| ||| | ||||| |||||
 Db 1522433 AACAAATGTATTCTATTGTTCAATCTAT 1522459


```
RESULT 15
US-09-822-830A-444/c
; Sequence 444, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Rehtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 444
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 535
; OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-444

Query Match          62.7%; Score 18.8; DB 9; Length 1630;
Best Local Similarity 76.7%; Pred. No. 2.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AACAAATGATGTCGGGTGATCATCTATGAC 30
    ||||| ||||| ||||| ||||| ||||| |||||
Db 633 AACGAAGTAAGCCCTGTGATCATCTATGAC 604

RESULT 16
US-10-027-632-8033/c
; Sequence 8033, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8033
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-8033

Query Match          62.0%; Score 18.6; DB 18; Length 665;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 21; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ACAATGATGTCGGGTGATCATCTATG 28
    ||||| ||||| ||||| ||||| ||||| |||||
Db 459 ACAATGATGTCGTAYGTGATGATG 433

RESULT 17
US-10-027-632-8033/c
; Sequence 8033, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8033
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-8033

Query Match          62.0%; Score 18.6; DB 18; Length 665;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 21; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ACAATGATGTCGGGTGATCATCTATG 28
    ||||| ||||| ||||| ||||| ||||| |||||
Db 459 ACAATGATGTCGTAYGTGATGATG 433

RESULT 18
US-10-372-095-3
; Sequence 3, Application US/10372095
; Publication No. US20030162256A1
; GENERAL INFORMATION:
; APPLICANT: Juppner, Harald
; APPLICANT: Rubin, David A.
; TITLE OF INVENTION: PTHrP and PTHrP Receptors, Methods and Uses Thereof
; FILE REFERENCE: 0609.4740002
; CURRENT APPLICATION NUMBER: US/10/372,095
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/449,632
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: US 60/110,467
; PRIOR FILING DATE: 1998-11-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2152
; TYPE: DNA
; ORGANISM: Danio rerio
```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (394)..(2019)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2125)..(2125)
; OTHER INFORMATION: n is any nucleotide of a,t,g or c
US-10-372-095-3

Query Match          62.0%; Score 18.6; DB 17; Length 2152;
Best Local Similarity 84.0%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 TGTATGTCGGGTGACATCTATGAC 30
Db 639 TCTGTGCTCTGAGTACATCTATGAC 663

RESULT 19
US-10-105-637-1
; Sequence 1, Application US/10105637
; Publication No. US20030087252A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; FILE REFERENCE: 529452500120
; CURRENT APPLICATION NUMBER: US/10/105,637
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 10/034,650
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 90442
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-105-637-1

Query Match          62.0%; Score 18.6; DB 15; Length 90442;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGACATCTATGAC 29
Db 3731 ATGTATGTCGGGSCACATATATGA 3755

RESULT 20
US-10-034-650-43
; Sequence 43, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 90442

```

```

; TYPE: DNA
; ORGANISM: Mus musculus
US-10-034-650-43

Query Match          62.0%; Score 18.6; DB 18; Length 90442;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGACATCTATGAC 29
Db 3731 ATGTATGTCGGGSCACATATATGA 3755

RESULT 21
US-10-424-599-132779/c
; Sequence 132779, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132779
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90906C.1
US-10-424-599-132779

Query Match          61.3%; Score 18.4; DB 19; Length 255;
Best Local Similarity 78.6%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATGAC 29
Db 105 ACTATGTTGTCGGTTCGACGCGATGA 78

RESULT 22
US-10-027-632-320106/c
; Sequence 320106, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 320106

```

```

; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-320106

Query Match      61.3%; Score 18.4; DB 14; Length 650;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCCGGTGTCATCTATGA 29
   | | | | | | | | | | | | | | | | | | | |
Db 260 AGAATATTGTCAGGTGTCATCTATGA 233

RESULT 23
US-10-027-632-320106/c
; Sequence 320106, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.139
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 320106
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-320106

Query Match      61.3%; Score 18.4; DB 18; Length 650;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCCGGTGTCATCTATGA 29
   | | | | | | | | | | | | | | | | | | | |
Db 260 AGAATATTGTCAGGTGTCATCTATGA 233

RESULT 24
US-09-925-065A-765698/c
; Sequence 765698, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 10:22:28 ; Search time 10140 Seconds
(without alignment)
120.778 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgatgtccgggtacatcatgac 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/PCRTUS1_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCRTUS2_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCRTUS3_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

```

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:*
118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:*
119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:*
120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:*
121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq:*
122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq:*
123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq:*
124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq:*
125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq:*
126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq:*
127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq:*
128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq:*
129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	35	US-09-806-197-24
2	30	100.0	2048	35	US-09-806-197-3
3	30	100.0	2048	35	US-09-806-197-4
4	30	100.0	2052	35	US-09-806-197-1
5	30	100.0	2052	35	US-09-806-197-7
6	30	100.0	2052	35	US-09-806-197-5
7	23.6	78.7	2046	20	US-09-300-487-2
8	21	70.0	15582	28	US-09-620-392-37594
9	21	70.0	15582	28	US-09-702-134-25342
10	21	70.0	15582	35	US-09-815-264-65330
11	20.6	68.7	214464	94	US-09-245-228-19
12	20.6	68.7	343544	116	US-60-466-412-86692
13	20.6	68.7	342222	67	US-10-990-328-95619
14	20.2	67.3	32768	90	US-60-213-178-293
15	20	66.7	4642	28	US-09-620-392-39645
16	20	66.7	4642	28	US-09-620-392-43697
17	20	66.7	4642	32	US-09-702-134-14042
18	20	66.7	4642	35	US-09-815-264-81213
19	20	66.7	20760	32	US-09-702-134-18180
20	20	66.7	20760	35	US-09-815-264-82022
21	20	66.7	21170	28	US-09-620-392-69029
22	19.6	65.3	380	29	US-09-654-617-396089
23	19.6	65.3	380	31	US-09-684-016-396089
24	19.6	65.3	380	54	US-10-437-963-26247
25	19.6	65.3	1729	2	PCT-US03-41761-60305
26	19.6	65.3	1729	2	PCT-US03-41768-60305
27	19.6	65.3	99231	35	US-09-803-736-1139
28	19.6	65.3	143879	35	US-09-803-736-1140
29	19.6	64.7	201	67	US-10-990-328-197684
30	19.4	64.7	2763	29	US-09-652-814-8994
31	19.4	64.7	2763	29	US-09-652-915-8287
32	19.4	64.7	67603	67	US-10-990-328-94217
33	19.4	64.7	374849	47	US-10-087-192-1627
34	19.4	64.0	656	89	US-60-207-458-26047
35	19.2	64.0	10450	28	US-09-620-392-25846
36	19.2	64.0	11263	28	US-09-620-392-45346
37	19.2	64.0	11263	32	US-09-702-134-8669
38	19.2	64.0	11263	35	US-09-815-264-75280
39	19.2	64.0	42772	47	US-10-087-192-1903
40	19.2	64.0	2731748	51	US-10-297-465A-1
41	19.2	64.0	2731748	51	US-10-297-465B-1
42	19.2	64.0	261	19	US-09-394-086A-4397
43	19	63.3	261	77	US-60-084-492-4322
44	19	63.3	484	26	US-09-594-596-1761
45	19	63.3	484	29	US-09-654-617-438086
46	19	63.3	484	31	US-09-684-016-438086
47	19	63.3	484	52	US-10-304-123A-1761
48	19	63.3	484	58	US-10-626-717-1761
49	19	63.3	512	45	US-10-021-323-6008
50	19	63.3			Sequence 6008, Ap

51	19	63.3	512	95	US-60-255-619-6008	Sequence 6008, Ap
52	19	63.3	1825	62	US-10-767-795-7186	Sequence 7186, Ap
53	19	63.3	2700	2	PCT-US03-03482-4157	Sequence 4157, Ap
54	19	63.3	2700	53	US-10-357-502-4157	Sequence 4157, Ap
55	19	63.3	5535	35	US-60-068-138-243	Sequence 243, Ap
56	19	63.3	35042	35	US-09-806-866A-2	Sequence 2, Appli
57	19	63.3	35042	65	US-10-915-740A-2	Sequence 2, Appli
58	19	63.3	50000	49	US-10-152-724A-22	Sequence 22, Appli
59	19	63.3	87039	35	US-09-803-736-1143	Sequence 1243, Ap
60	19	63.3	134402	35	US-09-803-736-1150	Sequence 1250, Ap
61	19	63.3	198632	66	US-10-940-774-12781	Sequence 12781, A
62	19	63.3	198637	66	US-10-940-774-17393	Sequence 17393, A
63	19	63.3	1596662	41	US-09-947-911-108	Sequence 108, App
64	19	63.3	1601042	28	US-09-634-106B-59064	Sequence 59064, A
65	19	63.3	1601042	46	US-10-027-632-59064	Sequence 59064, A
66	19	63.3	2242716	35	US-09-806-866A-1068	Sequence 1068, Ap
67	19	63.3	2242716	65	US-10-915-740A-1068	Sequence 1068, Ap
68	19	63.3	2272325	46	US-10-018-470A-1	Sequence 1, Appli
69	18.8	62.7	286	21	US-09-388-906A-765	Sequence 765, App
70	18.8	62.7	315	39	US-09-912-293-71909	Sequence 71909, A
71	18.8	62.7	323	27	US-09-606-977-11322	Sequence 11322, A
72	18.8	62.7	323	83	US-60-141-233-11322	Sequence 11322, A
73	18.8	62.7	365	26	US-09-565-306-26751	Sequence 26751, A
74	18.8	62.7	431	23	US-09-521-640-34199	Sequence 34199, A
75	18.8	62.7	431	26	US-09-552-087-11477	Sequence 11477, A
76	18.8	62.7	431	26	US-09-552-087B-11477	Sequence 11477, A
77	18.8	62.7	560	86	US-60-170-911-939	Sequence 939, App
78	18.8	62.7	1630	1	PCT-US01-10485A-444	Sequence 444, App
79	18.8	62.7	1630	35	US-09-823-830A-444	Sequence 444, App
80	18.8	62.7	46182	116	US-60-466-412-83125	Sequence 83125, A
81	18.8	62.7	52059	67	US-10-990-328-93411	Sequence 93411, A
82	18.6	62.0	56	2	PCT-US03-0871A-19775	Sequence 19775, A
83	18.6	62.0	56	3	PCT-US03-0491A-19775	Sequence 19775, A
84	18.6	62.0	201	116	US-60-461-762-15993	Sequence 15993, A
85	18.6	62.0	201	117	US-60-470-166-32767	Sequence 32767, A
86	18.6	62.0	242	25	US-09-540-212A-1214	Sequence 1214, Ap
87	18.6	62.0	368	23	US-09-521-640-174248	Sequence 174248, A
88	18.6	62.0	601	66	US-10-940-774-103581	Sequence 103581, A
89	18.6	62.0	665	28	US-09-634-306B-8033	Sequence 8033, Ap
90	18.6	62.0	665	46	US-10-027-633-8033	Sequence 28005, A
91	18.6	62.0	1611	51	US-10-266-090-28005	Sequence 3, Appli
92	18.6	62.0	2152	53	US-10-372-095-3	Sequence 552, App
93	18.6	62.0	4020	94	US-60-242-679-552	Sequence 1595, Ap
94	18.6	62.0	90442	1	PCT-US02-4141A-1595	Sequence 43, Appli
95	18.6	62.0	90442	46	US-10-034-650-43	Sequence 1595, Ap
96	18.6	62.0	90442	46	US-10-035-832-1595	Sequence 1, Appli
97	18.6	62.0	90442	48	US-10-105-637-1	Sequence 251, App
98	18.6	62.0	143552	48	US-10-144-779-251	Sequence 14608, A
99	18.6	62.0	256287	66	US-10-940-774-14608	Sequence 3414, Ap
100	18.6	62.0	381283	116	US-60-461-762-3414	

ALIGNMENTS

RESULT 1
 US-09-806-197-24
 ; Sequence 24, Application US/09806197
 ; GENERAL INFORMATION:
 ; APPLICANT: NISHIKAWA, SATOMI
 ; APPLICANT: OEDA, KENJI
 ; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
 ; FILE REFERENCE: 7372-70911
 ; CURRENT APPLICATION NUMBER: US/09/806,197
 ; CURRENT FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 24
 ; LENGTH: 30
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic

```
; OTHER INFORMATION: oligonucleotide
US-09-806-197-24
Query Match      100.0%; Score 30; DB 35; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
   |||||
Db 1 AACAAATGATGTCGGGTGTACATCTATGAC 30

RESULT 2
US-09-806-197-3
; Sequence 3, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806,197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-806-197-3

Query Match      100.0%; Score 30; DB 35; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
   |||||
Db 1742 AACAAATGATGTCGGGTGTACATCTATGAC 1771

RESULT 3
US-09-806-197-4
; Sequence 4, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806,197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-806-197-4

Query Match      100.0%; Score 30; DB 35; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
   |||||
Db 1742 AACAAATGATGTCGGGTGTACATCTATGAC 1771

RESULT 4
US-09-806-197-1
; Sequence 1, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
```

```
; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806,197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-806-197-1

Query Match      100.0%; Score 30; DB 35; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
   |||||
Db 1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775

RESULT 5
US-09-806-197-7
; Sequence 7, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806,197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-806-197-7

Query Match      100.0%; Score 30; DB 35; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
   |||||
Db 1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775

RESULT 6
US-09-806-197-5
; Sequence 5, Application US/09806197
; GENERAL INFORMATION:
; APPLICANT: NISHIKAWA, SATOMI
; APPLICANT: OEDA, KENJI
; TITLE OF INVENTION: PLANT PROMOTERS AND PLANT TERMINATORS
; FILE REFERENCE: 7372-70911
; CURRENT APPLICATION NUMBER: US/09/806,197
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2056
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-806-197-5

Query Match      100.0%; Score 30; DB 35; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
   |||||
Db 1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775
```

```
RESULT 7
US-09-300-487-2
; Sequence 2, Application US/09300487
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,434
; FILING DATE: 12-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1...2042
US-09-300-487-2
Query Match 78.7%; Score 23.6; DB 20; Length 2042;
Best Local Similarity 86.7%; Pred. No. 4.5;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
|||||
Db 1737 AACAAAGTTTGTCCGGTGTATATTATGAC 1766
|||||

RESULT 8
US-09-620-392-37594/c
; Sequence 37594, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15582
; TYPE: DNA
; ORGANISM: Oryza sativa
; INDIVIDUAL ISOLATE:
; FEATURE:
; NAME/KEY:
; LOCATION:
US-09-620-392-37594
Query Match 70.0%; Score 21; DB 28; Length 15582;
Best Local Similarity 82.8%; Pred. No. 1.1e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2 ACAATGATGTCGGGTGTACATCTATGAC 30
|||||
Db 1611 ACAATATTGTTGGTGTACATCTATGAC 1583
|||||

RESULT 9
US-09-702-134-25342/c
; Sequence 25342, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15582
; TYPE: DNA
; ORGANISM: Oryza sativa
; INDIVIDUAL ISOLATE:
; FEATURE:
; NAME/KEY:
; LOCATION:
US-09-702-134-25342
Query Match 70.0%; Score 21; DB 32; Length 15582;
Best Local Similarity 82.8%; Pred. No. 1.1e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2 ACAATGATGTCGGGTGTACATCTATGAC 30
|||||
Db 1611 ACAATATTGTTGGTGTACATCTATGAC 1583
|||||

RESULT 10
US-09-815-264-65330/c
; Sequence 65330, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15582
; TYPE: DNA
; ORGANISM: Oryza sativa
; INDIVIDUAL ISOLATE:
; FEATURE:
; NAME/KEY:
; LOCATION:
US-09-815-264-65330
Query Match 70.0%; Score 21; DB 35; Length 15582;
```

```
Best Local Similarity 82.8%; Pred. No. 1.1e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATG 30
Db 1611 ACAATATTTGTTGGTGACATGATGAC 1583

RESULT 11
US-60-245-228-19
; Sequence 19, Application US/60245228
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN CYCLASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN CYCLASE PROTEINS, AND USES
; FILE REFERENCE: CL000878
; CURRENT APPLICATION NUMBER: US/60/245,228
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 630
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 214464
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(214464)
; OTHER INFORMATION: n = A,T,C or G
US-60-245-228-19

Query Match 68.7%; Score 20.6; DB 94; Length 214464;
Best Local Similarity 85.2%; Pred. No. 3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATG 28
Db 46216 AAAACGTATGTCAGTGACATTTATG 46242

RESULT 12
US-60-466-412-86692
; Sequence 86692, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86692
; LENGTH: 343544
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-466-412-86692

Query Match 68.7%; Score 20.6; DB 116; Length 343544;
Best Local Similarity 85.2%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATG 28
Db 169241 AAAACGTATGTCAGTGACATTTATG 169267

RESULT 13
US-10-990-328-95619
; Sequence 95619, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95619
; LENGTH: 344222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(344222)
; OTHER INFORMATION: n = A,T,C or G
US-10-990-328-95619

Query Match 68.7%; Score 20.6; DB 67; Length 344222;
Best Local Similarity 85.2%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACAATGATGTCGGGTGACATCTATG 28
Db 169333 AAAACGTATGTCAGTGACATTTATG 169359

RESULT 14
US-60-213-178-293
; Sequence 293, Application US/60213178
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000689
; CURRENT APPLICATION NUMBER: US/60/213,178
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 1425
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 293
; LENGTH: 32768
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(32768)
; OTHER INFORMATION: n = A,T,C or G
US-60-213-178-293

Query Match 67.3%; Score 20.2; DB 90; Length 32768;
Best Local Similarity 88.0%; Pred. No. 3.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGACATCTATG 29
Db 4622 ATGTATGTCAGTGACATCTGTGA 4646

RESULT 15
US-09-620-392-39645
; Sequence 39645, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 39645
```



```
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-39645

Query Match      66.7%; Score 20; DB 28; Length 4642;
Best Local Similarity 82.1%; Pred. No. 2.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACAAATGTATGTCGGGTGTACATCTATG 28
Db 2252 AACAAATGTACTTCCCTTGTACATCTACG 2279

RESULT 16
US-09-620-392-43697
; Sequence 43697, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 43697
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-43697

Query Match      66.7%; Score 20; DB 28; Length 4642;
Best Local Similarity 82.1%; Pred. No. 2.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACAAATGTATGTCGGGTGTACATCTATG 28
Db 2252 AACAAATGTACTTCCCTTGTACATCTACG 2279

RESULT 17
US-09-702-134-14042
; Sequence 14042, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 14042
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-702-134-14042

Query Match      66.7%; Score 20; DB 32; Length 4642;
Best Local Similarity 82.1%; Pred. No. 2.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACAAATGTATGTCGGGTGTACATCTATG 28
Db 2252 AACAAATGTACTTCCCTTGTACATCTACG 2279

RESULT 18
US-09-815-264-81213
; Sequence 81213, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 81213
; LENGTH: 4642
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-81213

Query Match      66.7%; Score 20; DB 35; Length 4642;
Best Local Similarity 82.1%; Pred. No. 2.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACAAATGTATGTCGGGTGTACATCTATG 28
Db 2252 AACAAATGTACTTCCCTTGTACATCTACG 2279

RESULT 19
US-09-702-134-18180
; Sequence 18180, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 18180
; LENGTH: 20760
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-702-134-18180

Query Match      66.7%; Score 20; DB 32; Length 20760;
Best Local Similarity 82.1%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACAAATGTATGTCGGGTGTACATCTATG 28
Db 10899 AAAAATGCATCTTAGGTGTACATCTATG 10926

RESULT 20
US-09-815-264-82022
; Sequence 82022, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
```

```

/ AFFILIATION:  DRC, DRCongo
/ TITLE OF INVENTION:  38-21(15097)D
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER:  US/09/654,617
/ CURRENT FILING DATE:  2000-09-05
/ NUMBER OF SEQ ID NOS:  463173
/ SEQ ID NO 396089
/ LENGTH: 380
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ OTHER INFORMATION: unsure at all n locations

```

```

/ REFERENCE: 38-21(15097)D
/ TITLE OF INVENTION: Annotated Plant Genes
/ FILE REFERENCE: 38-21(15097)D
/ CURRENT APPLICATION NUMBER: US/09/654,617
/ CURRENT FILING DATE: 2000-09-05
/ NUMBER OF SEQ ID NOS: 463173
/ SEQ ID NO 396089
/ LENGTH: 380
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ OTHER INFORMATION: unsure at all n locations
/

```

```

; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(380)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31056C.1
; US-10-437-963-26247

```

```
Query Match          65.3%; Score 19.6; DB 54; Length 380;
Best Local Similarity 78.6%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

Qy 1 AACAAATGTATGTCGGGTGACATCTATG 28
Db 54 AAAAATGCATCTNNANGTGTACATCTATG 27

RESULT 25

```

PCT-US03-41761-60305
; Sequence 60305, Application PC/TUS0341761
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNise, Sue K.
; APPLICANT: CHARTERIS, Paul
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND SYSTEMS FOR INFERRING BOVINE BREED
; FILE REFERENCE: WM1150WO
; CURRENT APPLICATION NUMBER: PCT/US03/41761
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 60305
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Bovine
PCT-US03-41761-60305

```

Query Match	65.3%	Score 19.6;	DB 2;	Length 1729;
Best Local Similarity	84.6%	Pred. No. 3.4e+02;		
Matches 22; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy 4 AATGATATGCCGGTGTACATCTATGA 29
1592 AATGCATGTGTGGTGTACATATATGA 1617

Search completed: October 28, 2005, 14:07:33
Job time : 10149 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2005, 10:26:40 ; Search time 1143 Seconds
(without alignments)
377.374 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgtatgtccggtacatcatgac 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2723106 seqs, 7189968421 residues

Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Pending Patents NA New *

1:	/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2:	/cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
3:	/cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
4:	/cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
5:	/cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
6:	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
7:	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
8:	/cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
9:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
10:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
11:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
12:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
13:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
14:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
15:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
16:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
17:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
18:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
19:	/cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
20:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
21:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
22:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
23:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
24:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
25:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
26:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
27:	/cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
28:	/cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	2930	27	US-11-130-391-5 Sequence 5, Appli
2	20.6	68.7	344222	15	US-10-990-328A-95619 Sequence 95619, A
3	20.2	67.3	878	10	US-10-301-480C-141267 Sequence 141267, A
4	20.2	67.3	878	17	US-10-301-480A-141267 Sequence 141267, A
5	20.2	67.3	878	19	US-10-301-480B-141267 Sequence 141267, A

6	19.6	65.3	1114	7	US-09-925-065A-708892 Sequence 708892, A
7	19.6	65.3	1191	8	US-09-474-435A-147353 Sequence 147353, A
8	19.6	65.3	1230	8	US-09-474-435A-33673 Sequence 33673, A
9	19.6	65.3	1729	11	US-10-750-185-60305 Sequence 60305, A
10	19.6	65.3	1729	11	US-10-750-185-60305 Sequence 60305, A
11	19.6	65.3	4762	28	US-60-655-875-10887 Sequence 10887, A
12	19.4	64.7	201	15	US-10-990-328A-197684 Sequence 197684, A
13	19.4	64.7	551	7	US-09-925-065A-193679 Sequence 193679, A
14	19.4	64.7	561	10	US-10-301-480C-659145 Sequence 659145, A
15	19.4	64.7	561	17	US-10-301-480A-659145 Sequence 659145, A
16	19.4	64.7	561	18	US-10-301-480A-283076 Sequence 283076, A
17	19.4	64.7	561	18	US-10-301-480A-896485 Sequence 896485, A
18	19.4	64.7	561	19	US-10-301-480B-659145 Sequence 659145, A
19	19.4	64.7	67603	19	US-10-990-328A-94217 Sequence 94217, A
20	19	63.3	484	13	US-10-703-032-63131 Sequence 63131, A
21	19	63.3	543	10	US-10-301-480C-631688 Sequence 631688, A
22	19	63.3	543	17	US-10-301-480A-631688 Sequence 631688, A
23	19	63.3	543	18	US-10-301-480A-255619 Sequence 255619, A
24	19	63.3	543	18	US-10-301-480A-869028 Sequence 869028, A
25	19	63.3	543	19	US-10-301-480B-631688 Sequence 631688, A
26	19	63.3	545	7	US-09-925-065A-163111 Sequence 163111, A
27	19	63.3	551	7	US-09-925-065A-193680 Sequence 193680, A
28	19	63.3	561	10	US-10-301-480C-659146 Sequence 659146, A
29	19	63.3	561	17	US-10-301-480A-659146 Sequence 659146, A
30	19	63.3	561	18	US-10-301-480A-283077 Sequence 283077, A
31	19	63.3	561	18	US-10-301-480A-896486 Sequence 896486, A
32	19	63.3	561	19	US-10-301-480B-659146 Sequence 659146, A
33	19	63.3	3102	8	US-09-474-435A-142942 Sequence 142942, A
34	19	63.3	3393	8	US-09-474-435A-30235 Sequence 30235, A
35	19	63.3	198632	15	US-10-940-774A-12781 Sequence 12781, A
36	19	63.3	198637	15	US-10-940-774A-17393 Sequence 17393, A
37	18.8	62.7	575	1	PCT-US05-00517-1673 Sequence 1673, Ap
38	18.8	62.7	688	10	US-10-301-480C-242051 Sequence 242051, A
39	18.8	62.7	688	17	US-10-301-480A-242051 Sequence 242051, A
40	18.8	62.7	52099	15	US-10-990-328A-93411 Sequence 93411, A
41	18.6	62.0	601	15	US-10-940-774A-103581 Sequence 103581, A
42	18.6	62.0	256287	13	US-10-940-774A-14608 Sequence 14608, A
43	18.4	61.3	412	13	US-10-703-032-101063 Sequence 101063, A
44	18.4	61.3	658	7	US-09-925-065A-765698 Sequence 765698, A
45	18.4	61.3	993	10	US-10-301-480C-950134 Sequence 950134, A
46	18.4	61.3	993	17	US-10-301-480A-950134 Sequence 950134, A
47	18.4	61.3	993	18	US-10-301-480A-574065 Sequence 574065, A
48	18.4	61.3	993	18	US-10-301-480A-1187474 Sequence 1187474, A
49	18.4	61.3	993	19	US-10-301-480B-950134 Sequence 950134, A
50	18.4	61.3	1132	8	US-09-474-435A-58466 Sequence 58466, A
51	18.4	61.3	2501	9	US-10-548-548-26 Sequence 26, Appl
52	18.4	61.3	3177	24	US-11-097-143-40196 Sequence 40196, A
53	18.4	61.3	4354	24	US-11-097-143-13798 Sequence 13798, A
54	18.4	61.3	5130	8	US-09-474-435A-6445 Sequence 6445, Ap
55	18.4	61.3	5460	24	US-11-097-143-40195 Sequence 40195, A
56	18.4	61.3	40512	15	US-10-940-774A-16612 Sequence 16612, A
57	18.4	61.3	63118	1	PCT-US04-37982-2051 Sequence 2051, Ap
58	18.4	61.3	112613	15	US-10-990-328A-97760 Sequence 97760, A
59	18.4	61.3	242678	1	PCT-US04-37982-3211 Sequence 3211, Ap
60	18.2	60.7	4919	24	US-11-097-143-20047 Sequence 20047, A
61	18.2	60.7	539	9	US-10-550-533-12 Sequence 12, Appl
62	18	60.0	546	7	US-09-925-065A-620304 Sequence 620304, A
63	18	60.0	599	13	US-10-972-079-57689 Sequence 57689, A
64	18	60.0	599	13	US-09-925-065A-744005 Sequence 744005, A
65	18	60.0	703	7	US-10-301-480C-100411 Sequence 100411, A
66	18	60.0	722	10	US-10-301-480A-100411 Sequence 100411, A
67	18	60.0	722	17	US-10-301-480A-100411 Sequence 100411, A
68	18	60.0	722	19	US-10-301-480B-100411 Sequence 100411, A
69	18	60.0	751	10	US-10-301-480C-47989 Sequence 47989, A
70	18	60.0	751	17	US-10-301-480A-47989 Sequence 47989, A
71	18	60.0	751	17	US-10-301-480B-47989 Sequence 47989, A
72	18	60.0	784	10	US-10-301-480C-47988 Sequence 47988, A
73	18	60.0	784	17	US-10-301-480A-47988 Sequence 47988, A
74	18	60.0	784	19	US-10-301-480B-47988 Sequence 47988, A
75	18	60.0	997	17	US-10-301-480C-949724 Sequence 949724, A
76	18	60.0	997	17	US-10-301-480A-949724 Sequence 949724, A
77	18	60.0	997	18	US-10-301-480B-573655 Sequence 573655, A
78	18	60.0	997	18	US-10-301-480-1187064 Sequence 1187064, A

```
79      18      60.0      997      19      US-10-301-480B-949724      Sequence 949724.
c 80      18      60.0      6439      21      US-11-031-175-813      Sequence 813 App
81      18      60.0      154347      28      US-60-659-397-11390      Sequence 11390, A
c 82      17.8      59.3      64      27      US-11-130-645A-391878      Sequence 391878,
c 83      17.8      59.3      64      27      US-11-130-645A-409918      Sequence 409918,
c 84      17.8      59.3      64      27      US-11-130-645A-550200      Sequence 550200,
c 85      17.8      59.3      201      15      US-10-990-328A-504983      Sequence 504983,
c 86      17.8      59.3      201      15      US-10-990-328A-505004      Sequence 505004,
c 87      17.8      59.3      201      28      US-60-659-397-25908      Sequence 25908, A
c 88      17.8      59.3      201      28      US-60-659-397-26116      Sequence 26116, A
c 89      17.8      59.3      440      11      US-10-302-689A-116961      Sequence 116961,
c 90      17.8      59.3      440      12      US-10-450-763-2874      Sequence 2874, Ap
c 91      17.8      59.3      540      7      US-09-925-065A-222468      Sequence 222468,
c 92      17.8      59.3      540      7      US-09-925-065A-222469      Sequence 222469,
c 93      17.8      59.3      540      7      US-09-925-065A-222470      Sequence 222470,
c 94      17.8      59.3      541      10      US-10-301-480C-683633      Sequence 683633,
c 95      17.8      59.3      541      10      US-10-301-480C-683634      Sequence 683634,
c 96      17.8      59.3      541      10      US-10-301-480C-683635      Sequence 683635,
c 97      17.8      59.3      541      13      US-10-703-032-71753      Sequence 71753, A
c 98      17.8      59.3      541      17      US-10-301-480A-683633      Sequence 683633,
c 99      17.8      59.3      541      17      US-10-301-480A-683634      Sequence 683634,
c 100      17.8      59.3      541      17      US-10-301-480A-683635      Sequence 683635,

ALIGNMENTS

RESULT 1
US-11-130-391-5
; Sequence 5, Application US/11130391
; GENERAL INFORMATION:
; APPLICANT: Sumitomo Chemical Co., Ltd.
; TITLE OF INVENTION: HERBICIDAL COMPOUND RESISTANT PLANT
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/130,391
; PRIOR FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: JP 2004/147363
; PRIOR FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: JP 2005/070980
; PRIOR FILING DATE: 2005-03-14
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930)...(2561)
; FEATURE:
; OTHER INFORMATION: Part of an expression plasmid containing a Protoporphyrinogen IX
US-11-130-391-5
Query Match      100.0%; Score 30; DB 27; Length 2930;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      AACAAATGATGTCGGGTGATCATCTGAC 30
      |||
Db      615      AACAAATGATGTCGGGTGATCATCTGAC 644

RESULT 2
US-10-990-328A-95619
; Sequence 95619, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USING THEREOF
; FILE REFERENCE: CLO01495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95619
; LENGTH: 344222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(344222)
; OTHER INFORMATION: n = A,T,C or G
US-10-990-328A-95619
Query Match      68.7%; Score 20.6; DB 15; Length 344222;
Best Local Similarity 85.2%; Pred. No. 85;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      2      ACAATGATGTCGGGTGATCATCTATG 28
      |||
Db      169333      AAAACGATATGCCAGTGATCATTTATG 169359

RESULT 3
US-10-301-480C-141267
; Sequence 141267, Application US/10301480C
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480C
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 989478
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 141267
; LENGTH: 878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480C-141267
Query Match      67.3%; Score 20.2; DB 10; Length 878;
Best Local Similarity 88.0%; Pred. No. 70;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      5      ATGTATGTCGGGTGATCATCTATGA 29
      |||
Db      721      ATGTATGTCAGTGATCATCTGTGA 745

RESULT 4
US-10-301-480A-141267
; Sequence 141267, Application US/10301480A
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480A
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 989478
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 141267
; LENGTH: 878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480A-141267
```

Query Match 67.3%; Score 20.2; DB 17; Length 878;
 Best Local Similarity 88.0%; Pred. No. 70;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGTACATCTATGA 29
 Db 721 ATGTATGTCGAAGTGTACATCTGTGA 745

RESULT 5
 US-10-301-480B-141267
 ; Sequence 141267, Application US/10301480B
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; FILE REFERENCE: 108827-137
 ; CURRENT APPLICATION NUMBER: US/10/301,480B
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/215,598
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 989478
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 141267
 ; LENGTH: 878
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-301-480B-141267

Query Match 67.3%; Score 20.2; DB 19; Length 878;
 Best Local Similarity 88.0%; Pred. No. 70;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGTACATCTATGA 29
 Db 721 ATGTATGTCGAAGTGTACATCTGTGA 745

RESULT 6
 US-09-925-065A-708892
 ; Sequence 708892, Application US/09925065A
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; FILE REFERENCE: 108827.135
 ; CURRENT APPLICATION NUMBER: US/09/925,065A
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846
 ; PRIOR FILING DATE: 2001-05-09
 ; NUMBER OF SEQ ID NOS: 957086
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 708892
 ; LENGTH: 1114
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-925-065A-708892

Query Match 65.3%; Score 19.6; DB 7; Length 1114;
 Best Local Similarity 84.6%; Pred. No. 1.4e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AACAAATGTATGTCGGGTGTACATCTA 26

Db 214 AACAAATGTATGTCCTCTGTACCCCTA 239

RESULT 7
 US-09-474-435A-147353/c
 ; Sequence 147353, Application US/09474435A
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jingdong
 ; TITLE OF INVENTION: ARABIDOPSIS THALIANA GENOME SEQUENCE AND USES THEREOF
 ; FILE REFERENCE: 38-21(15473)B
 ; CURRENT APPLICATION NUMBER: US/09/474,435A
 ; CURRENT FILING DATE: 1999-12-28
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: US 60/089,524
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: US 60/089,516
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: US 60/089,808
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,812
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,807
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,806
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,811
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,813
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 09/333,534
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 60/111,990
 ; PRIOR FILING DATE: 1998-12-14
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 195836
 ; SEQ ID NO 147353
 ; LENGTH: 1191
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana landsberg
 US-09-474-435A-147353

Query Match 65.3%; Score 19.6; DB 8; Length 1191;
 Best Local Similarity 84.6%; Pred. No. 1.4e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 ATGTATGTCGGGTGTACATCTATGAC 30
 Db 985 AAGTACGTCCTCGTGTACATCTATGAC 960

RESULT 8
 US-09-474-435A-33673/c
 ; Sequence 33673, Application US/09474435A
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jingdong
 ; TITLE OF INVENTION: ARABIDOPSIS THALIANA GENOME SEQUENCE AND USES THEREOF
 ; FILE REFERENCE: 38-21(15473)B
 ; CURRENT APPLICATION NUMBER: US/09/474,435A
 ; CURRENT FILING DATE: 1999-12-28
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: US 60/089,524
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: US 60/089,516
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: US 60/089,808
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,812
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,807
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: US 60/089,806
 ; PRIOR FILING DATE: 1998-06-18

```
; PRIOR APPLICATION NUMBER: US 60/089,811
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,813
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 09/333,534
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 60/111,990
; PRIOR FILING DATE: 1998-12-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 195836
; SEQ ID NO 33673
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana landsberg
US-09-474-435A-33673

Query Match      65.3%; Score 19.6; DB 8; Length 1230;
Best Local Similarity 84.6%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 ATGTATGTCGGGTGACATCTATGAC 30
Db 985 AAGTACGTCTCGTGTACATCTATGAC 960

RESULT 9
US-10-750-185-60305
; Sequence 60305, Application US/10750185
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 60305
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Bovine 19866880911656
US-10-750-185-60305

Query Match      65.3%; Score 19.6; DB 11; Length 1729;
Best Local Similarity 84.6%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 AATGTATGTCGGGTGACATCTATGA 29
Db 1592 AATGCAIGTGTGTTGATACATATGA 1617

RESULT 10
US-10-623-60305
; Sequence 60305, Application US/10750623
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
```

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 60305
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Bovine 19866880911656
US-10-750-623-60305

Query Match      65.3%; Score 19.6; DB 11; Length 1729;
Best Local Similarity 84.6%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 AATGTATGTCGGGTGACATCTATGA 29
Db 1592 AATGCAIGTGTGTTGATACATATGA 1617

RESULT 11
US-60-655-875-10887
; Sequence 10887, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrei
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liliang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 10887
; LENGTH: 4762
; TYPE: DNA
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (324)..(324)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (326)..(326)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (342)..(342)
; OTHER INFORMATION: n is a, c, g, or t
US-60-655-875-10887

Query Match      65.3%; Score 19.6; DB 28; Length 4762;
Best Local Similarity 84.6%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CAATGTATGTCGGGTGACATCTATG 28
Db 4549 CAATGTATGTATTGTGTACATATG 4574

RESULT 12
US-10-990-328A-197684
; Sequence 197684, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
```

```

RESULT 14
US/10/301-480C-659145
; Sequence 659145, Application US/10301480C
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and M
; TITLE OF INVENTION: Nucleotide Polymorp
; FILE REFERENCE: 10827-137
; CURRENT APPLICATION NUMBER: US/10/301,48
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10

```

```
Query Match      64.7%; Score 19.4; DB 18; Length 561;
Best Local Similarity 79.3%; Pred. No. 1.6e+02;
Matches 23: Conservative 0; Mismatches 6; Indels 0; Gaps 0
```



```
Qy 1 AACAAATGATGTCGGGTGTACATCTATGA 29
    ||| ||||| ||||| ||||| |||||
Db 270 AACCTATATGTCGGGTGTCTTCTATGA 298

RESULT 17
US-10-301-480-896485
; Sequence 896485, Application US/10301480
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 896485
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-896485

Query Match 64.7%; Score 19.4; DB 18; Length 561;
Best Local Similarity 79.3%; Pred. No. 1.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGA 29
    ||| ||||| ||||| ||||| |||||
Db 270 AACCTATATGTCGGGTGTCTTCTATGA 298

RESULT 18
US-10-301-480B-659145
; Sequence 659145, Application US/10301480B
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480B
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 989478
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 659145
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480B-659145

Query Match 64.7%; Score 19.4; DB 19; Length 561;
Best Local Similarity 79.3%; Pred. No. 1.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGA 29
    ||| ||||| ||||| ||||| |||||
Db 270 AACCTATATGTCGGGTGTCTTCTATGA 298

RESULT 19
US-10-990-328A-94217
; Sequence 94217, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94217
; LENGTH: 67603
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328A-94217

Query Match 64.7%; Score 19.4; DB 15; Length 67603;
Best Local Similarity 79.3%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGA 29
    ||| ||||| ||||| ||||| |||||
Db 9084 AGCACTGCTTGACAGGTGTACATCTATGA 9112

RESULT 20
US-10-703-032-63131/c
; Sequence 63131, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Cao, Yongwei
; APPLICANT: Masucci, James D.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(S3374)B
; CURRENT APPLICATION NUMBER: US/10/703,032
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/020,338
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 211164
; SEQ ID NO 63131
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_TA_63131
US-10-703-032-63131

Query Match 63.3%; Score 19; DB 13; Length 484;
Best Local Similarity 81.5%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTAT 27
    ||| ||||| ||||| ||||| |||||
Db 223 AATGATATATGTGTGTGTACATCTAT 197

RESULT 21
US-10-301-480C-631688/c
; Sequence 631688, Application US/10301480C
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480C
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
```

; NUMBER OF SEQ ID NOS: 989478
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 631688
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480C-631688

Query Match 63.3%; Score 19; DB 10; Length 543;
Best Local Similarity 75.9%; Pred. No. 2.4e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGACATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 141 AACCCWATATGTCGTGTCCTTCTATGA 113

RESULT 22

US-10-301-480A-631688/c
; Sequence 631688, Application US/10301480A
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480A
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 989478
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 631688
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480A-631688

Query Match 63.3%; Score 19; DB 17; Length 543;
Best Local Similarity 75.9%; Pred. No. 2.4e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGACATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 141 AACCCWATATGTCGTGTCCTTCTATGA 113

RESULT 23

US-10-301-480-255619/c
; Sequence 255619, Application US/10301480
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 255619
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-255619

Query Match 63.3%; Score 19; DB 18; Length 543;
Best Local Similarity 75.9%; Pred. No. 2.4e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGACATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 141 AACCCWATATGTCGTGTCCTTCTATGA 113

RESULT 24

US-10-301-480-869028/c
; Sequence 869028, Application US/10301480
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 869028
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-869028

Query Match 63.3%; Score 19; DB 18; Length 543;
Best Local Similarity 75.9%; Pred. No. 2.4e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGACATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 141 AACCCWATATGTCGTGTCCTTCTATGA 113

RESULT 25

US-10-301-480B-631688/c
; Sequence 631688, Application US/10301480B
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827-137
; CURRENT APPLICATION NUMBER: US/10/301,480B
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 989478
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 631688
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480B-631688

Query Match 63.3%; Score 19; DB 19; Length 543;
Best Local Similarity 75.9%; Pred. No. 2.4e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGACATCTATGA 29
||| : ||||| ||||| ||||| |||||
Db 141 AACCCWATATGTCGTGTCCTTCTATGA 113

Search completed: October 28, 2005, 14:26:38
Job time : 1146 secs

THIS PAGE BLANK (USPTO)